GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 30, 2003, 15:23:05 ; Search time 39 Seconds
(without alignments)
675.647 Million cell updates/sec

Title: US-09-806-536A-14
Perfect score: 1386
Sequence: 1 MAAAEPSPRRVGFVGAGRMA.....AATMSAVEAATCRAKELSRK 274

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	613	44.2	279	2	T23765	hypothetical prote
2	584.5	42.2	319	1	A41770	pyrroline-5-carbox
3	502.5	36.3	273	2	T06477	probable pyrroline
4	492	35.5	274	1	S10186	pyrroline-5-carbox
5	480	34.6	270	2	G97299	pyrroline-5-carbox
6	475	34.3	276	1	JQ2334	pyrroline-5-carbox
7	469	33.8	266	2	AE1124	1-pyrroline-5-carb
8	458	33.0	266	2	AG1484	1-pyrroline-5-carb
9	455.5	32.9	270	2	AG1867	pyrroline-5-carbox
10	423.5	30.6	269	1	RDECC	pyrroline-5-carbox
11	423.5	30.6	269	2	D90683	pyrroline-5-carbox
12	423.5	30.6	269	2	H85533	pyrroline-5-carbox
13	420.5	30.3	284	2	T36286	pyrroline-5-carbox
14	419.5	30.3	269	2	AI0549	pyrroline-5-carbox
15	412.5	29.8	267	2	S76767	probable pyrroline
16	395	28.5	265	2	H95107	pyrroline-5-carbox
17	390.5	28.2	263	2	D71281	probable pyrroline
18	387	27.9	261	2	JC2078	pyrroline-5-carbox
19	385	27.8	265	2	B97976	pyrroline-5-carbox
20	383.5	27.7	294	2	S72897	pyrroline-5-carbox
21	366	26.4	295	2	G70745	probable proC prot
22	362	26.1	299	2	T29226	hypothetical prote
23	360.5	26.0	277	2	H90194	hypothetical prote
24	354	25.5	270	2	E69682	pyrroline-5-carbox
25	354	25.5	279	2	G83837	pyrroline-5-carbox
26	349.5	25.2	264	2	C75385	pyrroline-5-carbox
27	349	25.2	262	2	D86860	pyrroline-5-carbox
28	345.5	24.9	273	2	JQ0418	pyrroline-5-carbox
29	344	24.8	282	2	T50305	Delta 1-pyrroline-
30	337.5	24.4	278	1	G69964	pyrroline-5-carbox
31	335	24.2	266	2	G72769	probable pyrroline
32	322	23.2	305	2	C82524	pyrroline-5-carbox
33	321	23.2	286	2	S25293	pyrroline-5-carbox
34	318.5	23.0	265	2	F70315	pyrroline carboxyl
35	312.5	22.5	267	2	G83760	pyrroline-5-carbox
36	310.5	22.4	311	2	S57863	pyrroline-5-carbox
37	305.5	22.0	320	2	JC4830	pyrroline-5-carbox
38	301.5	21.8	271	2	I64060	pyrroline-5-carbox
39	300.5	21.7	272	2	AH2847	pyrroline-5-carbox
40	300.5	21.7	274	2	G97624	delta 1-pyrroline-
41	298	21.5	272	2	D82321	pyrroline-5-carbox
42	297.5	21.5	255	2	E72360	pyrroline-5-carbox
43	290	20.9	260	2	E87310	pyrroline-5-carbox
44	283.5	20.5	300	2	AD3573	pyrroline-5-carbox
45	276	19.9	271	2	G89929	hypothetical prote

ALIGNMENTS

RESULT 1
T23765
hypothetical protein M153.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T23765
R;Matthews, P.
submitted to the EMBL Data Library, November 1995
A;Reference number: Z19796
A;Accession: T23765
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-279 <WIL>
A;Cross-references: EMBL:Z67995; PIDN:CAA91943.1; GSPDB:GN00028; CESP:M153.1
A;Experimental source: clone M153
C;Genetics:
A;Gene: CESP:M153.1
A;Map position: X
A;Introns: 23/1; 106/3; 177/3; 248/3
C;Superfamily: pyrroline-5-carboxylate reductase

Query Match 44.2%; Score 613; DB 2; Length 279;
Best Local Similarity 46.1%; Pred. No. 7.9e-41;
Matches 123; Conservative 56; Mismatches 86; Indels 2; Gaps 1;

```
Qy         10 RVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAP--TDRNLCHFQALGCRTTHSNQEVLQ 67
              ::||:|||:|| |:|:||| :|:: | :|:||:|   :  |   :|||  ||| | ||:|
Db          2 KIGFIGAGKMAQALARGLINSGRITADNIIASSPKRDEVFLDQCKALGLNTTHDNAEVVQ 61

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNL 127
                 :|  | ||  :  | :|:|| :: ||::||:| |:::  :| |||  :||:||:||
Db         62 KSDVVFLAVKPVHVSKVASEIAPALSKEHLVVSIALGITIRNIESLLPTKSRVVRVMPNT 121

Qy        128 PCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAF 187
              | ||: ||   | |      : : :: ||   |   |||| ::|  |||||||  ::: |
Db        122 PSVVRAGASAFAMGSACRDGDAETVEKLLSTVGFAVEVPEIHIDPVTGLSGSGPSYMFAV 181

Qy        188 SEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHA 247
               | ||:| ||:|:|  || ::|| |||| |||:|  | |||||: || :| |:::||:|
Db        182 IEGLADGGVKVGLPRDLALKLAAYTLLGAAKMVLETGIHPAQLKDDVQSPAGSSVYGMHK 241

Qy        248 LEQGGLRAATMSAVEAATCRAKELSRK 274
              || |||:   | |||||| |::    |
Db        242 LESGGLKGVLMDAVEAATNRSRATGDK 268
```

RESULT 2
A41770
pyrroline-5-carboxylate reductase (EC 1.5.1.2) - human
N;Alternate names: P5C reductase
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

```
C;Accession: A41770
R;Dougherty, K.M.; Brandriss, M.C.; Valle, D.
J. Biol. Chem. 267, 871-875, 1992
A;Title: Cloning human pyrroline-5-carboxylate reductase cDNA by complementation in Sacc
A;Reference number: A41770; MUID:92112821; PMID:1730675
A;Accession: A41770
A;Molecule type: mRNA
A;Residues: 1-319 <DOU>
A;Cross-references: GB:M77836; NID:g189497; PIDN:AAA36407.1; PID:g189498
A;Note: sequence extracted from NCBI backbone (NCBIN:75606, NCBIP:75608)
C;Genetics:
A;Gene: GDB:PYCR1; P5C; PYCR
A;Cross-references: GDB:135716; OMIM:179035
A;Map position: 17pter-17qter
C;Superfamily: pyrroline-5-carboxylate reductase
C;Keywords: oxidoreductase; proline biosynthesis

  Query Match             42.2%;  Score 584.5;  DB 1;  Length 319;
  Best Local Similarity   46.4%;  Pred. No. 1.6e-38;
  Matches  124;  Conservative   47;  Mismatches   89;  Indels    7;  Gaps    3;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
Db          3 VGFIGAGQLAFALAKGFTAAGVLAAHKIMASSP-DMDLATVSALRKMGVKLTPHNKETVQ 61

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLP---PNTRVLRVL 124
Db         62 HSDVLFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRPAPRVIRCM 121

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
Db        122 TNTPVVVREGATVYATGTHAQVEDGRLMEQLLSTVGFCTEVEEDLIDAVTGLSGSGPAYA 181

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
Db        182 FTALDALADGGVKMGLPRRLAVRLGAQALLGAAKMLLHSEQHPGQLKDNVSSPGGATIHA 241

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
Db        242 LHVLESGGFRSLLINAVEASCIRTREL 268


RESULT 3
T06477
probable pyrroline-5-carboxylate reductase (EC 1.5.1.2) - garden pea
C;Species: Pisum sativum (garden pea)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C;Accession: T06477
R;Williamson, C.L.; Slocum, R.D.
Plant Physiol. 100, 1464-1470, 1992
A;Title: Molecular cloning and evidence for osmoregulation of the d1-pyrroline-5-carboxy
A;Reference number: Z15706
A;Accession: T06477
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-273 <WIL>
A;Cross-references: EMBL:X62842; NID:g20850; PIDN:CAA44646.1; PID:g20851
A;Experimental source: cv. Wando
C;Genetics:
A;Gene: proC
C;Superfamily: pyrroline-5-carboxylate reductase
C;Keywords: oxidoreductase; proline biosynthesis

  Query Match             36.3%;  Score 502.5;  DB 2;  Length 273;
  Best Local Similarity   39.7%;  Pred. No. 3.6e-32;
  Matches  104;  Conservative   58;  Mismatches   99;  Indels    1;  Gaps    1;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTHSNQEVLQSCL 70
Db         13 LGFIGAGKMAESIAKGASRSGVLPSSRIVTAHSNPSRRAAFESIGITVLSSNDDVVRASN 72

Qy         71 LVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNLPCV 130
Db         73 VVVFSVKPQLVKDVVLKLKPLLTKDKLLVSVAAGIKLKDLQE-WAGHERFIRVMPNTPAA 131

Qy        131 VQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAFSEA 190
Db        132 VGQAASVMSLGGAATEEDANLISQLFGSIGKIWKADDKFFDAITGLSGSGPAYIYLAIEA 191

Qy        191 LAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHALEQ 250
Db        192 LADGGVAAGLPRDLALSLASQTVLGAASMATLSGKHPGQLKDDVTSPGGTTIAGVHELEK 251

Qy        251 GGLRAATMSAVEAATCRAKELS 272
Db        252 GGFRGTLMNAVVAAAKRSRELS 273


RESULT 4
S10186
pyrroline-5-carboxylate reductase (EC 1.5.1.2) - soybean
C;Species: Glycine max (soybean)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S10186
R;Delauney, A.J.; Verma, D.P.S.
Mol. Gen. Genet. 221, 299-305, 1990
A;Title: A soybean gene encoding Delta1-pyrroline-5-carboxylate reductase was isolate
A;Reference number: S10186; MUID:90340278; PMID:2199815
A;Accession: S10186
A;Molecule type: mRNA
A;Residues: 1-274 <DEL>
A;Cross-references: EMBL:X16352; NID:g18723; PIDN:CAA34401.1; PID:g18724
A;Note: the authors translated the codon GAA for residue 115 as Pro and GTG for resid
C;Superfamily: pyrroline-5-carboxylate reductase
C;Keywords: oxidoreductase; proline biosynthesis

  Query Match             35.5%;  Score 492;  DB 1;  Length 274;
  Best Local Similarity   39.6%;  Pred. No. 2.4e-31;
  Matches  107;  Conservative   51;  Mismatches   96;  Indels   16;  Gaps    3;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCH--------FQALGCRTTHSN 62
Db         13 LGFIGAGKMAESIARGAVRSGVLPPSRI-------RTAVHFNLARRGAFESFGVTVLPSN 65

Qy         63 QEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLR 122
Db         66 DDVVRESDVVVLSVKPQLVKDVVSKLTPLLTKHKLLVSVAAGTKLKDLQE-WAGNDRFIR 124

Qy        123 VLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVA 182
Db        125 VMPNTPAAVGQAASVMSLGGSATEEDGNIIAQLFGSIGKIWKAEEKYFDAITGLSGSGPA 184

Qy        183 FVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTI 242
Db        185 YVYLAIEALADGGVAAGLPRDLSLSLASQTVLGAASMVSQTGKHPGQLKDDVTSPGGTTI 244

Qy        243 YGLHALEQGGLRAATMSAVEAATCRAKELS 272
Db        245 TGIHELENGGFRGTLMNAVVAAAKRSRELS 274


RESULT 5
G97299
pyrroline-5-carboxylate reductase [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C;Accession: G97299
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: G97299
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-270 <KUR>
```

```
A;Cross-references: GB:AE001437; PIDN:AAK81186.1; PID:g15026326; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC3252
C;Superfamily: pyrroline-5-carboxylate reductase

  Query Match             34.6%;  Score 480;  DB 2;  Length 270;
  Best Local Similarity   37.5%;  Pred. No. 2.1e-30;
  Matches  100;  Conservative   64;  Mismatches  101;  Indels    2;  Gaps    2;

Qy         10 RVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHF-QALGCRTTHSNQEVLQS 68
              :|||:| | || || :|:::|  |  ::|: | |::  |    :  |  ||: |::|
Db          4 KVGFIGCGNMAQAIIKGMVKAKVVPNENIIVSNPSNEKLEKISKECGVLTTNDNKQVALK 63

Qy         69 CLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLP-PNTRVLRVLPNL 127
                 ::: : ||:    |::|:  :|    |:|::|||||:     :    | ||:| :||
Db         64 ADIIVLSVKPNKYEKVISEIKDLVQQSVIIVAIAAGVSIEKTRIMFKNSNIRVVRAMPNT 123

Qy        128 PCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAF 187
              | :| |    ::    :   | | :  : |: |:|| | || ::  ||:|||  |:|  |
Db        124 PALVGEAMTAISPCEEIDKGELKNVTEIFESFGKCEVVDEALMNAVTGVSGSSPAYVYMF 183

Qy        188 SEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHA 247
              ||:|: ||  ||    |:: |||:|| |||:|  |:|| :|: |||:| ||||  ::|
Db        184 VEAMADAAVLNGMTREKAYKFAAQSVLGAAKMILETGEHPGKLKDDVCSPSGTTIEAVYA 243

Qy        248 LEQGGLRAATMSAVEAATCRAKELSRK 274
              ||: | ||: ::|||:|   ::| :| |
Db        244 LEKSGFRASVIAAVDACIKKSKLMSSK 270


RESULT 6
JQ2334
pyrroline-5-carboxylate reductase (EC 1.5.1.2) [similarity] - Arabidopsis thaliana
N;Alternate names: protein T9L3_100
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Sep-2000
C;Accession: JQ2334; T51425
R;Verbruggen, N.; Villarroel, R.; Van Montagu, M.
Plant Physiol. 103, 771-781, 1993
A;Title: Osmoregulation of a pyrroline-5-carboxylate reductase gene in Arabidopsis thali
A;Reference number: JQ2334; MUID:94294559; PMID:8022935
A;Accession: JQ2334
A;Molecule type: DNA
A;Residues: 1-276 <VER>
A;Cross-references: GB:M76538; NID:g166814; PIDN:AAA61346.1; PID:g166815
R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A;Reference number: Z25394
A;Accession: T51425
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-276 <SAT>
A;Cross-references: EMBL:AL391149
A;Experimental source: cultivar Columbia; BAC clone T9L3
C;Comment: This enzyme is involved in the last step of the proline biosynthetic pathway.
C;Genetics:
A;Map position: 5
A;Introns: 67/3; 84/1; 114/3; 139/1; 179/2; 214/3
A;Note: T9L3_100
C;Superfamily: pyrroline-5-carboxylate reductase
C;Keywords: oxidoreductase; proline biosynthesis

  Query Match             34.3%;  Score 475;  DB 1;  Length 276;
  Best Local Similarity   40.0%;  Pred. No. 5.3e-30;
  Matches  106;  Conservative   54;  Mismatches  103;  Indels    2;  Gaps    2;

Qy         10 RVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCH-FQALGCRTTHSNQEVLQS 68
              :|||:|||:|| :|||:|: :| :    |  :  :: |    |:: |     :::||::
Db         12 KVGFIGAGKMAESIARGVVASGVLPPNRICTAVHSNLNRRDVFESFGVNVFSTSEEVVKE 71

Qy         69 CLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNLP 128
                 :|||: || |:   : |:   ::   ||||||||: |: |:|      | :||:|| |
Db         72 SDVVIFSVKPQVVKKAVTELKSKLSKNKILVSVAAGIKLNDLQE-WSGQDRFIRVMPNTP 130

Qy        129 CVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAFS 188
                | |' | ||: |      :  ::  |  | |:  :  |   |  |||||| |::
Db        131 AAVGEAASVMSLGTGATEEDGAIVAMLFGAVGKILKADEKMFDAVTGLSGSGPAYIFLAI 190

Qy        189 EALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHAL 248
              ||||:| |  |:|  ||  :|:||:|| | |:   |:||  |: || :||||||  |:| |
Db        191 EALADGGVAAGLPRELALSLASQTVLGAATMVSKTGKHPGVLKDDVTSPGGTTIAGVHEL 250

Qy        249 EQGGLRAATMSAVEAATCRAKELSR 273
              |:|  ||  |:|| ||  |::|||:
Db        251 EKGSFRATLMNAVVAAAKRSRELSQ 275


RESULT 7
AE1124
1-pyrroline-5-carboxylate reductase (ProC) homolog lmo0396 [imported] - Listeria mono
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AE1124
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AE1124
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-266 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAC98475.1; PID:g16409774; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo0396
C;Superfamily: pyrroline-5-carboxylate reductase

  Query Match             33.8%;  Score 469;  DB 2;  Length 266;
  Best Local Similarity   38.8%;  Pred. No. 1.5e-29;
  Matches  101;  Conservative   55;  Mismatches   96;  Indels    8;  Gaps    2;

Qy         10 RVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL-----GCRTTHSNQE 64
              |:||:||| |  |: :|| :|  |: : |:       |||   : |     | : |   ::
Db          3 RIGFIGAGNMGTAMIRGLAQANLVKREEIIVGG---RNLEKLKPLEAEFTGLQITTDTEK 59

Qy         65 VLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVL 124
              :::   ::| | ||: :| :|  |    :| : |::||||||::  ||||    |:::||:
Db         60 LVEQADIIILAVKPYTIPEILTSVKEKLTPDKIIISVAAGVTIQDLEELTSAKTKIVRVM 119

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
              || | :| |    ::   :| | | | :  :  : |  | | |  :|   |:|||  |:|
Db        120 PNTPALVGEAMSSVSPNTNVTSEELKEVTAIFTSFGEAEVVSENLMDAVIGVSGSSPAYV 179

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
                | ||||:|||  |||   |:: ||| :|| || :|  |:|| :|:  | :||||||
Db        180 YMFIEALADGAVLSGMPRDKAYKFAAQAVLGAAKTVLETGEHPGKLKDMVTSPGGTTIEA 239

Qy        245 LHALEQGGLRAATMSAVEAA 264
              : :||  | |:| ::||:||
Db        240 VKSLENDGFRSAVINAVQAA 259


RESULT 8
AG1484
1-pyrroline-5-carboxylate reductase (ProC) homolog lin0414 [imported] - Listeria inno
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AG1484
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
D.; Jones, L.M.; Karst, U.
```

```
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AG1484
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-266 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC95647.1; PID:g16412843; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin0414
C;Superfamily: pyrroline-5-carboxylate reductase

  Query Match             33.0%;  Score 458;  DB 2;  Length 266;
  Best Local Similarity   37.5%;  Pred. No. 1.1e-28;
  Matches   98;  Conservative   58;  Mismatches   97;  Indels    8;  Gaps    2;

Qy          9 RRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL-----GCRTTHSNQ 63
              :::||:|||  |  |: :|| :|   : : ||      |::   : |     | : |
Db          2 KKIGFIGAGNMGAAMIRGLAKANLGKEEDILVCG---RDMEKLKPLKTEFNGIQLTTDIA 58

Qy         64 EVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRV 123
              :: :   ::| : ||: :| :|  |   :| | |::||||||::  ||||    |:::||
Db         59 QLAEQADIIILSVKPYTIPEILTAVKDKITPEKIVISVAAGVTIKDLEELTSTETKIVRV 118

Qy        124 LPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAF 183
              :|| | :| |    ::   :| : ||  :  :  : |: | | |  :|   |:|||  |:
Db        119 MPNTPALVGEAMSSISPNSNVTAEETAHISTIFSSFGKAEVVAENLMDAVVGVSGSSPAY 178

Qy        184 VCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIY 243
              |  | ||||:|||  |||   |:: ||| :|| |||:|  |:|| :|:  | :||||||
Db        179 VYMFIEALADGAVLKGMPRDKAYKFAAQAVLGAAKMVLETGEHPGKLKDMVTSPGGTTIE 238

Qy        244 GLHALEQGGLRAATMSAVEAA 264
               : :||  | |:: :|||:||
Db        239 AVKSLEDTGFRSSVISAVQAA 259


RESULT 9
AG1867
pyrroline-5-carboxylate reductase [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AG1867
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AG1867
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-270 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB72446.1; PID:g17129833; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr0488
C;Superfamily: pyrroline-5-carboxylate reductase

  Query Match             32.9%;  Score 455.5;  DB 2;  Length 270;
  Best Local Similarity   40.9%;  Pred. No. 1.7e-28;
  Matches  110;  Conservative   43;  Mismatches  109;  Indels    7;  Gaps    6;

Qy         10 RVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAP-TDRNLCHFQALGCRTTHSNQEVLQS 68
              : | :| | |  |:   ||  |  :   :: | |- | |      |       |  | |
Db          4 KFGLIGGGVMGEALLSRLIARGIYQPSEVIVSEPQTARQAFLQQKYHVGVTTDNSLVFTQ 63

Qy         69 CLLVIF-ATKPHVLPAVLAEVAPVVTTEH--ILVSVAAGVSLSTLEELLPPNTRVLRVLP 125
                 |:| | || |  |:  |:|  | |:|  ::||: ||||||  ||    | : |:| :|
Db         64 AQDVVFLAVKPQVFSAIAQELADTVFTDHSPLVVSILAGVSLSQLEAAF-PQSPVIRAMP 122

Qy        126 NLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVC 185
              | |  |  |   :  | :  :   || | :  | |   || |: :|  |||||||  |:|
Db        123 NTPATVGSGMTAICSGAYTYAQHQKLAQQIFSAVGEVVEVSESLMDAVTGLSGSGPAYVA 182

Qy        186 AFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQ-HPAQLRSDVCTPGGTTIYG 244
                 ||||:|  |  |:|  :|:::| ||:||||: |||| : |||:|:  | :||||||
Db        183 LLVEALADGGVASGLPRGIANQLALQTVLGTAQ-LLHESKLHPAELKDRVTSPGGTTIAG 241

Qy        245 LHALEQGGLRAATMSAVEAATCRAKELSR 273
               :  ||: | |:| : ||:||| |::|| :
Db        242 IAQLEKAGFRSALIEAVKAATWRSQELGK 270


RESULT 10
RDECC
pyrroline-5-carboxylate reductase (EC 1.5.1.2) - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 01-Mar-2002
C;Accession: A00385; B64767
R;Deutch, A.H.; Smith, C.J.; Rushlow, K.E.; Kretschmer, P.J.
Nucleic Acids Res. 10, 7701-7714, 1982
A;Title: Escherichia coli delta(1)-pyrroline-5-carboxylate reductase: gene sequence,
A;Reference number: A00385; MUID:83116986; PMID:6296787
A;Accession: A00385
A;Molecule type: DNA
A;Residues: 1-269 <DEU>
A;Cross-references: GB:J01665; NID:g147358; PIDN:AAA86433.1; PID:g147359
A;Note: parts of this sequence, including the amino and carboxyl ends of the mature p
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: B64767
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-269 <BLAT>
A;Cross-references: GB:AE000145; GB:U00096; NID:g1786580; PIDN:AAC73489.1; PID:g17865
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: proC
A;Map position: 9 min
C;Function:
A;Description: catalyzes reduction of pyrroline-5-carboxylate to proline
A;Pathway: proline biosynthesis
A;Note: third enzyme in the proline biosynthetic pathway
C;Superfamily: pyrroline-5-carboxylate reductase
C;Keywords: oxidoreductase; proline biosynthesis

  Query Match             30.6%;  Score 423.5;  DB 1;  Length 269;
  Best Local Similarity   34.6%;  Pred. No. 5.6e-26;
  Matches   92;  Conservative   56;  Mismatches  117;  Indels    1;  Gaps    1;

Qy          9 RRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPT-DRNLCHFQALGCRTTHSNQEVLQ 67
              :::||:| | |  ||  ||| :|:|    |    |: |:        |     | ||| |
Db          3 KKIGFIGCGNMGKAILGGLIASGQVLPGQIWVYTPSPDKVAALHDQFGINAAESAQEVAQ 62

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNL 127
                 ::  | || ::  ||:|:   :  : ::||:|||||:|  |   |  : :::| :||
Db         63 IADIIFAAVKPGIMIKVLSEITSSLNKDSLVVSIAAGVTLDQLARALGHDRKIIRAMPNT 122

Qy        128 PCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAF 187
              | :|  |   :     |  :|  : ::    |  | : | :   |:|||  |:|  |
Db        123 PALVNAGMTSVTPNALVTPEDTADVLNIFRCFGEAEVIAEPMIHPVVGVSGSSPAYVFMF 182

Qy        188 SEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHA 247
              ||:|: ||  ||| : |:: ||| ::|:||||:|  |:||  |:  ||:|||||  :
Db        183 IEAMADAAVLGGMPRAQAYKFAAQAVMGSAKMVLETGEHPGALKDMVCSPGGTTIEAVRV 242

Qy        248 LEQGGLRAATMSAVEAATCRAKELSR 273
              ||: | ||| : |:     ::::||:
Db        243 LEEKGFRAAVIEAMTKCMEKSEKLSK 268
```

```
RESULT 11
D90683
pyrroline-5-carboxylate reductase [imported] - Escherichia coli (strain O157:H7, substra
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: D90683
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: D90683
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-269 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB33859.1; PID:g13359893; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs0437
C;Superfamily: pyrroline-5-carboxylate reductase

  Query Match             30.6%;  Score 423.5;  DB 2;  Length 269;
  Best Local Similarity   34.6%;  Pred. No. 5.6e-26;
  Matches   92;  Conservative   56;  Mismatches  117;  Indels    1;  Gaps    1;

Qy          9 RRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPT-DRNLCHFQALGCRTTHSNQEVLQ 67
              :::||:| | |  ||  ||| :|:|    |    |: |:       `|     | ||| |
Db          3 KKIGFIGCGNMGKAILGGLIASGQVLPGQIWVYTPSPDKVAALHDKFGINAAESAQEVAQ 62

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNL 127
                ::  | || ::  ||:|:   :  : ::||:||||:|  |   |  : :::| :||
Db         63 IADIIFAAVKPGIMIKVLSEITSSLNKDSLVVSIAAGVTLDQLARALGHDRKIIRAMPNT 122

Qy        128 PCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAF 187
              | :|  |   :     |  :|  : ::    |  | : |  :    |:|||  |:|  |
Db        123 PALVNAGMTSVTPNALVTPEDTADVLNIFRCFGEAEVIAEPMIHPVVGVSGSSPAYVFMF 182

Qy        188 SEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHA 247
               ||:|: ||  ||| : |:: ||| ::|:|||:|  |:||  |:  ||:|||||  :
Db        183 IEAMADAAVLGGMPRAQAYKFAAQAVMGSAKMVLETGEHPGALKDMVCSPGGTTIEAVRV 242

Qy        248 LEQGGLRAATMSAVEAATCRAKELSR 273
              ||: | ||| : |:     ::::||:
Db        243 LEEKGFRAAVIEAMTKCMEKSEKLSK 268


RESULT 12
H85533
pyrroline-5-carboxylate reductase [imported] - Escherichia coli (strain O157:H7, substra
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: H85533
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: H85533
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-269 <STO>
A;Cross-references: GB:AE005174; NID:g12513226; PIDN:AAG54732.1; GSPDB:GN00145; UWGP:Z04
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: proC
C;Superfamily: pyrroline-5-carboxylate reductase

  Query Match             30.6%;  Score 423.5;  DB 2;  Length 269;
  Best Local Similarity   34.6%;  Pred. No. 5.6e-26;
  Matches   92;  Conservative   56;  Mismatches  117;  Indels    1;  Gaps    1;

Qy          9 RRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPT-DRNLCHFQALGCRTTHSNQEVLQ 67
              :::||:| | |  ||  ||| :|:|    |    |: |:       `|     | ||| |
Db          3 KKIGFIGCGNMGKAILGGLIASGQVLPGQIWVYTPSPDKVAALHDKFGINAAESAQEVAQ 62

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNL 127
                ::  | || ::  ||:|:   :  : ::||:||||:|  |   |  : :::| :||
Db         63 IADIIFAAVKPGIMIKVLSEITSSLNKDSLVVSIAAGVTLDQLARALGHDRKIIRAMPNT 122

Qy        128 PCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAF 187
              | :|  |   :     |  :|  : ::    |  | : |  :    |:|||  |:|  |
Db        123 PALVNAGMTSVTPNALVTPEDTADVLNIFRCFGEAEVIAEPMIHPVVGVSGSSPAYVFMF 182

Qy        188 SEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHA 247
               ||:|: ||  ||| : |:: ||| ::|:|||:|  |:||  |:  ||:|||||  :
Db        183 IEAMADAAVLGGMPRAQAYKFAAQAVMGSAKMVLETGEHPGALKDMVCSPGGTTIEAVRV 242

Qy        248 LEQGGLRAATMSAVEAATCRAKELSR 273
              ||: | ||| : |:     ::::||:
Db        243 LEEKGFRAAVIEAMTKCMEKSEKLSK 268


RESULT 13
T36286
pyrroline-5-carboxylate reductase - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C;Accession: T36286
R;Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M
submitted to the EMBL Data Library, May 1998
A;Reference number: Z21603
A;Accession: T36286
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-284 <SEE>
A;Cross-references: EMBL:AL049819; PIDN:CAB42663.1; GSPDB:GN00070; SCOEDB:SCE7.04c
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: proC; SCOEDB:SCE7.04c
C;Superfamily: pyrroline-5-carboxylate reductase

  Query Match             30.3%;  Score 420.5;  DB 2;  Length 284;
  Best Local Similarity   35.4%;  Pred. No. 1e-25;
  Matches   95;  Conservative   53;  Mismatches  119;  Indels    1;  Gaps    1;

Qy          5 EPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTHSNQE 64
              ||  ::|  :| |::  |:  |:| ||   |  :| :|         :|    |  :| |
Db         13 EPMTQKVAVLGTGKIGEALLSGMIGAGWAPAD-LLVTARRRERADELRARHGVTPVTNAE 71

Qy         65 VLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVL 124
                ::   :|   ||  :  :| |:|| |  : :::| ||||  |  || | | | |:||:
Db         72 AAKAADTLILTVKPQDMGTLLDELAPHVPADRLVISGAAGVPTSFFEERLAPGTPVVRVM 131

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
              | | :| |   |:: | |  ::      : :  | |:   |||:  |  | ||||| |:
Db        132 TNTPALVDEAMSVISAGTHATAAHLTHTEEIFGAVGKTLRVPESQQDACTALSGSGPAYF 191

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
                ||: :  : :|:|   || :  |: :| ||||   |:|| :|| :| :| ||||
Db        192 FYLVEAMTDAGILLGLPRDKAHDLIVQSAIGAAKMLRDSGEHPVKLRENVTSPAGTTINA 251

Qy        245 LHALEQGGLRAATMSAVEAATCRAKELS 272
              : ,||  |:||| ::|:|||  |::||:
Db        252 IRELENHGVRAALIAALEAARDRSRELA 279


RESULT 14
AI0549
pyrroline-5-carboxylate reductase [imported] - Salmonella enterica subsp. enterica se
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AI0549
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
```

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AI0549
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-269 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD08841.1; PID:g16501655; GSPDB:GN00176
C;Genetics:
A;Gene: STY0419
C;Superfamily: pyrroline-5-carboxylate reductase

Query Match 30.3%; Score 419.5; DB 2; Length 269;
Best Local Similarity 34.6%; Pred. No. 1.2e-25;
Matches 92; Conservative 57; Mismatches 116; Indels 1; Gaps 1;

```
Qy          9 RRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPT-DRNLCHFQALGCRTTHSNQEVLQ 67
              :::||:| | |  ||  ||| :|:|    |    |: |:        |     | ||| |
Db          3 KKIGFIGCGNMGKAILGGLIASGQVLPGQIWVYTPSPDKVAALHDQYGINAAQSAQEVAQ 62

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNL 127
                :|  | ||  ::  ||:|::  :  : ::||:||||:|  |   |  : :::| :||
Db         63 VADIVFGAVKPGIMVKVLSEISSSLNKDSLVVSIAAGVTLDQLARALGHDRKIIRAMPNT 122

Qy        128 PCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAF 187
              | :|  |   :     |   :|  : ::    |  | : |  :    |:|||  |:|  |
Db        123 PSLVNAGMTSVTPNALVTPEDTADVLNIFRCFGEAEVIAEPMIHPVVGVSGSSPAYVFMF 182

Qy        188 SEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHA 247
               ||:|: ||  ||| : |:: ||| ::|:||::|  |:|| :|:  ||:||||||  :
Db        183 IEAMADAAVLGGMPRAQAYKFAAQAVMGSAKIVLETGKHPGELKDMVCSPGGTTIEAVRV 242

Qy        248 LEQGGLRAATMSAVEAATCRAKELSR 273
              ||: | ||| : |:     ::: ||:
Db        243 LEERGFRAAVIEAMTKCMEKSEALSK 268
```

RESULT 15
S76767
probable pyrroline-5-carboxylate reductase (EC 1.5.1.2) - Synechocystis sp. (strain PCC
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S76767
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76767
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-267 <KAN>
A;Cross-references: EMBL:D90916; GB:AB001339; NID:g1653715; PIDN:BAA18679.1; PID:g165376
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Genetics:
A;Start codon: GTG
C;Superfamily: pyrroline-5-carboxylate reductase
C;Keywords: oxidoreductase; proline biosynthesis

Query Match 29.8%; Score 412.5; DB 2; Length 267;
Best Local Similarity 36.2%; Pred. No. 4.1e-25;
Matches 98; Conservative 50; Mismatches 110; Indels 13; Gaps 4;

```
Qy         10 RVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQ-ALGCRTTHSNQEVLQS 68
              ::| :| | || ||   ||       : |:   |      : |     | :  |||
Db          4 QLGIIGGGVMAEAILARLIAEKTYAPEEIIVGEPHGARRDYLQKTYQVRVSPDNQEAANV 63

Qy         69 CLLVIFATKPHVLPAVLAEVA-----PVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRV 123
                :::  | || ||  ||| :|     |:|      :|: |||||  :::  | :  ::|
Db         64 SEVLLLAVKPQVLDRVLASLAGGANRPLV------ISILAGVSLQRIQKGFPDHA-IIRA 116

Qy        124 LPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAF 183
              :|| |  |  |   :|  : |   :    : :  | |   ||||  :|  ||:|||| |:
Db        117 MPNTPATVGAGMTAIAANKMVEPDQLAKAKAIFSAVGNVVEVPENLMDAVTGVSGSGPAY 176

Qy        184 VCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIY 243
              |    ||||:| |  |:| ::| ::| ||:||||:::    :||||::  | :||||||
Db        177 VALMIEALADGGVLAGLPRAIAQKLALQTVLGTAELIKETEEHPAQIKDKVTSPGGTTIA 236

Qy        244 GLHALEQGGLRAATMSAVEAATCRAKELSRK 274
              |:  ||: | |:| : || ||  |::|| :|
Db        237 GVAVLEKMGFRSAIIEAVRAAYRRSQELGKK 267
```

Search completed: July 30, 2003, 15:28:05
Job time : 40 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 30, 2003, 15:23:57 ; Search time 22 Seconds
(without alignments)
1479.099 Million cell updates/sec

Title: US-09-806-536A-14
Perfect score: 1386
Sequence: 1 MAAAEPSPRRVGFVGAGRMA.....AATMSAVEAATCRAKELSRK 274

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1386	100.0	274	15	US-10-161-418A-11	Sequence 11, Appl
2	1386	100.0	274	15	US-10-161-418A-13	Sequence 13, Appl
3	584.5	42.2	319	15	US-10-161-418A-10	Sequence 10, Appl
4	584.5	42.2	319	15	US-10-161-418A-12	Sequence 12, Appl
5	560.5	40.4	315	9	US-09-912-717-3	Sequence 3, Appli
6	537.5	38.8	314	9	US-09-912-717-1	Sequence 1, Appli
7	477	34.4	319	15	US-10-161-418A-14	Sequence 14, Appl
8	467.5	33.7	255	10	US-09-925-300-1218	Sequence 1218, Ap
9	408.5	29.5	269	15	US-10-156-761-12258	Sequence 12258, A
10	401	28.9	284	15	US-10-128-714-3252	Sequence 3252, Ap
11	401	28.9	284	15	US-10-128-714-8252	Sequence 8252, Ap
12	366	26.4	295	10	US-09-712-363-174	Sequence 174, App
13	362.5	26.2	270	10	US-09-738-626-3960	Sequence 3960, Ap
14	116	8.4	144	9	US-09-939-980-445	Sequence 445, App
15	114	8.2	299	9	US-09-912-020-256	Sequence 256, App
16	95.5	6.9	549	15	US-10-156-761-14029	Sequence 14029, A
17	95	6.9	807	11	US-09-930-020A-2	Sequence 2, Appli
18	93.5	6.7	289	15	US-10-156-761-9583	Sequence 9583, Ap
19	93	6.7	2436	9	US-09-795-693-8	Sequence 8, Appli
20	93	6.7	2436	15	US-10-156-239-8	Sequence 8, Appli
21	93	6.7	2436	15	US-10-199-485-8	Sequence 8, Appli
22	92	6.6	408	10	US-09-712-363-273	Sequence 273, App
23	91.5	6.6	452	15	US-10-156-761-12493	Sequence 12493, A
24	89.5	6.5	323	15	US-10-156-761-9372	Sequence 9372, Ap
25	89.5	6.5	336	15	US-10-156-761-10217	Sequence 10217, A
26	89.5	6.5	488	15	US-10-156-761-7949	Sequence 7949, Ap
27	89.5	6.5	858	9	US-09-815-242-11396	Sequence 11396, A
28	89	6.4	447	15	US-10-156-761-9880	Sequence 9880, Ap
29	87.5	6.3	1024	15	US-10-211-962-44	Sequence 44, Appl
30	87.5	6.3	6145	15	US-10-156-761-7962	Sequence 7962, Ap
31	87	6.3	426	15	US-10-156-761-8581	Sequence 8581, Ap
32	87	6.3	949	9	US-09-841-835-10	Sequence 10, Appl
33	87	6.3	1327	9	US-09-841-835-2	Sequence 2, Appli
34	87	6.3	1327	11	US-09-972-115A-8	Sequence 8, Appli
35	86.5	6.2	340	9	US-09-815-242-10852	Sequence 10852, A
36	86.5	6.2	342	9	US-09-815-242-4924	Sequence 4924, Ap
37	86.5	6.2	1835	15	US-10-156-761-7963	Sequence 7963, Ap
38	86.5	6.2	4809	15	US-10-156-761-9090	Sequence 9090, Ap
39	86	6.2	492	9	US-09-815-242-11079	Sequence 11079, A
40	86	6.2	499	15	US-10-156-761-9271	Sequence 9271, Ap
41	85.5	6.2	342	15	US-10-156-761-14291	Sequence 14291, A
42	85.5	6.2	748	9	US-09-815-242-12792	Sequence 12792, A
43	85.5	6.2	792	9	US-09-815-242-12327	Sequence 12327, A
44	85.5	6.2	1687	14	US-10-094-679-3	Sequence 3, Appli
45	85	6.1	1074	9	US-09-509-196A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-161-418A-11
; Sequence 11, Application US/10161418A
; Publication No. US20030036078A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: P5CRs AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-055
; CURRENT APPLICATION NUMBER: US/10/161,418A
; CURRENT FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 60/296,080
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,509
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-161-418A-11

Query Match 100.0%; Score 1386; DB 15; Length 274;
Best Local Similarity 100.0%; Pred. No. 1.4e-134;
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60

Qy         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120

Qy        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180
```

```
Qy        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240

Qy        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274
              ||||||||||||||||||||||||||||||||||
Db        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274


RESULT 2
US-10-161-418A-13
; Sequence 13, Application US/10161418A
; Publication No. US20030036078A1
; GENERAL INFORMATION:
;  APPLICANT: EXELIXIS, INC.
;  TITLE OF INVENTION: P5CRs AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
;  FILE REFERENCE: EX02-055
;  CURRENT APPLICATION NUMBER: US/10/161,418A
;  CURRENT FILING DATE:  2002-09-11
;  PRIOR APPLICATION NUMBER: US 60/296,080
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/328,509
;  PRIOR FILING DATE: 2001-10-10
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
;   LENGTH: 274
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-161-418A-13

  Query Match             100.0%;  Score 1386;  DB 15;  Length 274;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-134;
  Matches  274;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60

Qy         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120

Qy        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180

Qy        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240

Qy        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274
              ||||||||||||||||||||||||||||||||||
Db        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274


RESULT 3
US-10-161-418A-10
; Sequence 10, Application US/10161418A
; Publication No. US20030036078A1
; GENERAL INFORMATION:
;  APPLICANT: EXELIXIS, INC.
;  TITLE OF INVENTION: P5CRs AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
;  FILE REFERENCE: EX02-055
;  CURRENT APPLICATION NUMBER: US/10/161,418A
;  CURRENT FILING DATE:  2002-09-11
;  PRIOR APPLICATION NUMBER: US 60/296,080
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/328,509
;  PRIOR FILING DATE: 2001-10-10
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
;   LENGTH: 319
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-161-418A-10

  Query Match             42.2%;  Score 584.5;  DB 15;  Length 319;
  Best Local Similarity   46.4%;  Pred. No. 6.6e-52;
  Matches  124;  Conservative   47;  Mismatches   89;  Indels    7;  Gaps    3;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
              |||:|||::|.|:|:|   || : |  |:||:| | :|    ||   | : |  |:| :|
Db          3 VGFIGAGQLAFALAKGFTAAGVLAAHKIMASSP-DMDLATVSALRKMGVKLTPHNKETVQ 61

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLP---PNTRVLRVL 124
                 ::  | |||::| :| |:   :   ||:|| ||||::|::|: |    | ||:| :
Db         62 HSDVLFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRPAPRVIRCM 121

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
               | | ||:||| | | | |    : :|:: ||   | | || |  :|  |||||||  |:
Db        122 TNTPVVVREGATVYATGTHAQVEDGRLMEQLLSTVGFCTEVEEDLIDAVTGLSGSGPAYA 181

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
                :|||:| ||||:|  || |: || ||| ||||||  ||| ||: :| :||| ||:
Db        182 FTALDALADGGVKMGLPRRLAVRLGAQALLGAAKMLLHSEQHPGQLKDNVSSPGGATIHA 241

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
              || || || |:  ::||||:  | :||
Db        242 LHVLESGGFRSLLINAVEASCIRTREL 268


RESULT 4
US-10-161-418A-12
; Sequence 12, Application US/10161418A
; Publication No. US20030036078A1
; GENERAL INFORMATION:
;  APPLICANT: EXELIXIS, INC.
;  TITLE OF INVENTION: P5CRs AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
;  FILE REFERENCE: EX02-055
;  CURRENT APPLICATION NUMBER: US/10/161,418A
;  CURRENT FILING DATE:  2002-09-11
;  PRIOR APPLICATION NUMBER: US 60/296,080
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/328,509
;  PRIOR FILING DATE: 2001-10-10
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
;   LENGTH: 319
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-161-418A-12

  Query Match             42.2%;  Score 584.5;  DB 15;  Length 319;
  Best Local Similarity   46.4%;  Pred. No. 6.6e-52;
  Matches  124;  Conservative   47;  Mismatches   89;  Indels    7;  Gaps    3;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
              |||:|||::|.|:|:|   || : |  |:||:| | :|    ||   | : |  |:| :|
Db          3 VGFIGAGQLAFALAKGFTAAGVLAAHKIMASSP-DMDLATVSALRKMGVKLTPHNKETVQ 61

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLP---PNTRVLRVL 124
                 ::  | |||::| :| |:   :   ||:|| ||||::|::|: |    | ||:| :
Db         62 HSDVLFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRPAPRVIRCM 121

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
               | | ||:||| | | | |    : :|:: ||   | | || |  :|  |||||||  |:
Db        122 TNTPVVVREGATVYATGTHAQVEDGRLMEQLLSTVGFCTEVEEDLIDAVTGLSGSGPAYA 181

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
                :|||:| ||||:|  || |: || ||| ||||||  ||| ||: :| :||| ||:
Db        182 FTALDALADGGVKMGLPRRLAVRLGAQALLGAAKMLLHSEQHPGQLKDNVSSPGGATIHA 241

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
```

```
                ||  ||  ||  |:  ::||||:  | :||
Db        242 LHVLESGGFRSLLINAVEASCIRTREL 268


RESULT 5
US-09-912-717-3
; Sequence 3, Application US/09912717
; Patent No. US20020081691A1
;   GENERAL INFORMATION:
;        APPLICANT: Hillman, Jennifer L.
;                   Corley, Neil C.
;                   Baughn, Mariah R.
;        TITLE OF INVENTION: DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE
;                            HOMOLOG
;        NUMBER OF SEQUENCES: 3
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Incyte Pharmaceuticals, Inc.
;             STREET: 3174 Porter Drive
;             CITY: Palo Alto
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: Windows
;             SOFTWARE: FastSEQ for Windows Version 2.0b
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/912,717
;             FILING DATE: 24-Jul-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/565,910
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Cerrone, Michael C
;             REGISTRATION NUMBER: 39,132
;             REFERENCE/DOCKET NUMBER: PF-0532 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650-855-0555
;             TELEFAX: 650-855-0572
;             TELEX: <Unknown>
;    INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 315 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: GENBANK
;             CLONE: 189498
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-912-717-3

  Query Match             40.4%;  Score 560.5;  DB 9;  Length 315;
  Best Local Similarity   46.1%;  Pred. No. 1.9e-49;
  Matches  123;  Conservative   46;  Mismatches   87;  Indels   11;  Gaps    5;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
              |||:|||::  | |:|   || : |  |:||:| | :|    ||   | : |  |:| :|
Db          3 VGFIGAGQL--AFAKGFTAAGVLAAHKIMASSP-DMDLATVSALRKMGVKLTPHNKETVQ 59

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLP---PNTRVLRVL 124
                ::  | |||::| :| |:   :   ||:|| ||||::|::|: |    |  ||:| :
Db         60 HSDVLFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRPAPRVIRCM 119

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
              | | ||:||| | | |    : :|:: ||   | | || |  :|  |||||| |:
Db        120 TNTPVVVREGATVYATGTHAQVEDGRLMEQLLSTVGFCTEVEEDLIDAVTGLSGSGPAY- 178

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
                ||:   |:| ||||:|  || |: || ||| ||||||  ||| ||: :| :||| ||:
Db        179 -AFTALDADGGVKMGLPRRLAVRLGAQALLGAAKMLLHSEQHPGQLKDNVSSPGGATIHA 237

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
                ||  ||  ||  |:  ::||||:  | :||
Db        238 LHVLESGGFRSLLINAVEASCIRTREL 264


RESULT 6
US-09-912-717-1
; Sequence 1, Application US/09912717
; Patent No. US20020081691A1
;   GENERAL INFORMATION:
;        APPLICANT: Hillman, Jennifer L.
;                   Corley, Neil C.
;                   Baughn, Mariah R.
;        TITLE OF INVENTION: DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE
;                            HOMOLOG
;        NUMBER OF SEQUENCES: 3
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Incyte Pharmaceuticals, Inc.
;             STREET: 3174 Porter Drive
;             CITY: Palo Alto
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: Windows
;             SOFTWARE: FastSEQ for Windows Version 2.0b
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/912,717
;             FILING DATE: 24-Jul-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/565,910
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Cerrone, Michael C
;             REGISTRATION NUMBER: 39,132
;             REFERENCE/DOCKET NUMBER:.PF-0532 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650-855-0555
;             TELEFAX: 650-855-0572
;             TELEX: <Unknown>
;    INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 314 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: PROSNON01
;             CLONE: 2278458
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-912-717-1

  Query Match             38.8%;  Score 537.5;  DB 9;  Length 314;
  Best Local Similarity   45.3%;  Pred. No. 4.5e-47;
  Matches  121;  Conservative   45;  Mismatches   88;  Indels   13;  Gaps    5;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
              |||:|||:     :|     || : |  |:||:| : ||    ||   |   | ||:| ::
Db          3 VGFIGAGQ----LAYRFTAAGILSAHKIIASSP-EMNLPTVSALRKMGVNLTRSNKETVK 57

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELL---PPNTRVLRVL 124
                ::  | |||::| :| |:   |   ||:|| ||||::|::|: |    |  :|:| :
Db         58 HSDVLFLAVKPHIIPFILDEIGADVQARHIVVSCAAGVTISSVEKKLMAFQPAPKVIRCM 117

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
              | | ||||||  | | | |    : :||: |: : | | || |  :|  |||||| |:
Db        118 TNTPVVVQEGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGSGPAY- 176

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
```

```
                 ||    |:| ||||:|  || :: || ||| |||||   ||| ||: :||:||| ||:
Db        177 -AFMALDADGGVKMGLPRRLAIQLGAQALLGAAKMLLDSEQHPCQLKDNVCSPGGATIHA 235

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
              || || || |:  ::||||: .| :||
Db        236 LHFLESGGFRSLLINAVEASCIRTREL 262


RESULT 7
US-10-161-418A-14
; Sequence 14, Application US/10161418A
; Publication No. US20030036078A1
; GENERAL INFORMATION:
;  APPLICANT: EXELIXIS, INC.
;  TITLE OF INVENTION: P5CRs AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
;  FILE REFERENCE: EX02-055
;  CURRENT APPLICATION NUMBER: US/10/161,418A
;  CURRENT FILING DATE:  2002-09-11
;  PRIOR APPLICATION NUMBER: US 60/296,080
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/328,509
;  PRIOR FILING DATE: 2001-10-10
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
;   LENGTH: 319
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-161-418A-14

  Query Match             34.4%;  Score 477;  DB 15;  Length 319;
  Best Local Similarity   41.0%;  Pred. No. 7.9e-41;
  Matches  110;  Conservative   47;  Mismatches  101;  Indels   10;  Gaps    4;

Qy         11 VGFVGAGRMA----GAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTHSNQEVL 66
              |||:|||::|    || ::  |  |  :   :  : |   :    |  ||      |
Db          3 VGFIGAGQLAMLWRGA-SRPQIPVGSQDNSQLPRNEPA--HGVRAQEDGCEPDTQQQGDG 59

Qy         67 QSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELL---PPNTRVLRV 123
              ::    :   : |::| :| |:   |   ||:|| ||||::|::|: |    |  :|:|
Db         60 EAQRRPVSGCEAHIIPFILVEIGADVQARHIVVSCAAGVTISSVEKKLMEFQPAPKVIRC 119

Qy        124 LPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAF 183
              : | | |||||| | | | |    : :||: |: : | | || |  :|  ||||||  |:
Db        120 MTNTPVVVQEGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGSKPAY 179

Qy        184 VCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIY 243
                  :|||:| ||||:|  || :: || ||| |||||   ||| ||: :||:||| ||:
Db        180 AFMALDALADGGVKMGLPRRLAIQLGAQALLGAAKMLLDSEQHPCQLKDNVCSPGGATIH 239

Qy        244 GLHALEQGGLRAATMSAVEAATCRAKEL 271
              || || || |:  ::||||:  | :||
Db        240 ALHFLESGGFRSLLINAVEASCIRTREL 267


RESULT 8
US-09-925-300-1218
; Sequence 1218, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
;  APPLICANT: Craig Rosen,
;  APPLICANT:  Steve Ruben
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA101
;  CURRENT APPLICATION NUMBER: US/09/925,300
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05988
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 1890
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1218
;   LENGTH: 255
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (2)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1218

  Query Match             33.7%;  Score 467.5;  DB 10;  Length 255;
  Best Local Similarity   48.5%;  Pred. No. 5.5e-40;
  Matches   95;  Conservative   34;  Mismatches   64;  Indels    3;  Gaps    1;

Qy         79 HVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLP---PNTRVLRVLPNLPCVVQEGA 135
              |::| :| |:   :   ||:|| |||||::|::|: |    |  ||:| : | | ||:|||
Db          9 HIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRPAPRVIRCMTNTPVVVREGA 68

Qy        136 IVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAFSEALAEGA 195
              | | | |    : :|:: || : | | || |  :|  |||||||  |:     :|||:|
Db         69 TVYATGTHAQVEDGRLMEQLLSSVGFCTEVEEDLIDAVTGLSGSGPAYAFTALDALADGG 128

Qy        196 VKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHALEQGGLRA 255
              ||||:|  || |: || ||| ||||||  ||| ||: :| :||| ||: || || || |:
Db        129 VKMGLPRRLAVRLGAQALLGAAKMLLHSEQHPGQLKDNVSSPGGATIHALHVLESGGFRS 188

Qy        256 ATMSAVEAATCRAKEL 271
                ::||||:  | :||
Db        189 LLINAVEASCIRTREL 204


RESULT 9
US-10-156-761-12258
; Sequence 12258, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12258
;   LENGTH: 269
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-12258

  Query Match             29.5%;  Score 408.5;  DB 15;  Length 269;
  Best Local Similarity   35.3%;  Pred. No. 7.1e-34;
  Matches   96;  Conservative   51;  Mismatches  108;  Indels   17;  Gaps    3;

Qy          9 RRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTD-----RNLCHFQALGCR---TTH 60
              ::|  :| |::  |:  |:|||          ||||     |     : |  |   |
Db          3 QKVAVLGTGKIGEALLSGMIRAG---------WAPTDLLVTARRPERAEELRARHGVTPV 53

Qy         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120
              :| |  ::   :|   ||  :  :| |:|| | |: :::| |||:  | ||| |   | |
Db         54 TNAEAAKTADTLILTVKPQDMGTLLTELAPHVPTDRLVISGAAGIPTSFLEERLAQGTPV 113

Qy        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180
              :||: | | :| |   |:: | |  :       : :  | |:   |||:  |  | |||||
Db        114 VRVMTNTPALVDEAMSVISAGSHATAGHLAHAEEIFGAVGKTLRVPESQQDACTALSGSG 173
```

```
Qy        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240
              |:     ||: :  : :|:|   || :  |: :| | ||   |:|| :|| :| :| ||
Db        174 PAYFFYLVEAMTDAGILLGLPRDKAHDLIVQSAIGAAIMLRDSGEHPVKLRENVTSPAGT 233

Qy        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELS 272
              ||  :  ||  |:||| ::|:||| |::||:
Db        234 TINAIRELENHGVRAALIAALEAARDRSRELA 265


RESULT 10
US-10-128-714-3252
; Sequence 3252, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Bo
;  APPLICANT:  Hu, Wenqi
;  APPLICANT:  Tishkoff, Daniel
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Eroshkin, Alexey M
;  APPLICANT:  Lemieux, Sebastien M
;  TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
;  TITLE OF INVENTION:  Methods of Use
;  FILE REFERENCE: 10182-018-999
;  CURRENT APPLICATION NUMBER: US/10/128,714
;  CURRENT FILING DATE:  2002-04-23
;  PRIOR APPLICATION NUMBER: US 60/285,697
;  PRIOR FILING DATE: 2001-04-23
;  PRIOR APPLICATION NUMBER: US 60/287,066
;  PRIOR FILING DATE: 2001-04-27
;  PRIOR APPLICATION NUMBER: US 60/295,890
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/303,899
;  PRIOR FILING DATE: 2001-07-09
;  PRIOR APPLICATION NUMBER: US 60/316,362
;  PRIOR FILING DATE: 2001-08-31
;  NUMBER OF SEQ ID NOS: 8603
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3252
;   LENGTH: 284
;   TYPE: PRT
;   ORGANISM: Aspergillus fumigatus
US-10-128-714-3252

  Query Match             28.9%;  Score 401;  DB 15;  Length 284;
  Best Local Similarity   37.3%;  Pred. No. 4.5e-33;
  Matches  101;  Conservative   45;  Mismatches  115;  Indels   10;  Gaps    4;

Qy         10 RVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTHSNQEVLQSC 69
              :: |:| | || ||  ||: :  :   :|  | | | |      || :|| || |   :
Db          8 KLAFIGGGNMASAIIGGLV-SQDMNPANITVSEPWDVNREKIAKLGVQTTTSNGEAAANA 66

Qy         70 LLVIFATKPHVLPAVLAEVAPV---VTTEHILVSVAAGVSLSTLEELLPPN----TRVLR 122
               :|| | ||     |  |:|      |:  ::||:|||::|::::| |  |      ::|
Db         67 DIVIIAVKPQTTKNVCQELATAWSQRTSLPVVVSIAAGITLNSMKEWLRTNDGRTAHIVR 126

Qy        123 VLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEE--VPEAYVDIHTGLSGSG 180
              |:|| | :|:|||  :     | : | :|:  ||::  :  |     |  :|: |||||||
Db        127 VMPNTPALVKEGASGLLASDDVTAEEKELIGALLQSVSKATEWVEKEELLDVVTGLSGSG 186

Qy        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240
               |:  |  | |   | :|:    | |:|||| ||  |||:   : ||||| :| :| ||
Db        187 PAYFFAMVEHLVASATALGLSKEQATRLAAQTCLGAGKMLVESSEEPAQLRKNVTSPNGT 246

Qy        241 TIYGLHALEQGGLRAATMSAVEAATCRAKEL 271
              |   |   |    :     ||:||| || ||
Db        247 THAALQTFESLNFKEIVDKAVQAATSRAAEL 277


RESULT 11
US-10-128-714-8252
; Sequence 8252, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Bo
;  APPLICANT:  Hu, Wenqi
;  APPLICANT:  Tishkoff, Daniel
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Eroshkin, Alexey M
;  APPLICANT:  Lemieux, Sebastien M
;  TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
;  TITLE OF INVENTION:  Methods of Use
;  FILE REFERENCE: 10182-018-999
;  CURRENT APPLICATION NUMBER: US/10/128,714
;  CURRENT FILING DATE:  2002-04-23
;  PRIOR APPLICATION NUMBER: US 60/285,697
;  PRIOR FILING DATE: 2001-04-23
;  PRIOR APPLICATION NUMBER: US 60/287,066
;  PRIOR FILING DATE: 2001-04-27
;  PRIOR APPLICATION NUMBER: US 60/295,890
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/303,899
;  PRIOR FILING DATE: 2001-07-09
;  PRIOR APPLICATION NUMBER: US 60/316,362
;  PRIOR FILING DATE: 2001-08-31
;  NUMBER OF SEQ ID NOS: 8603
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8252
;   LENGTH: 284
;   TYPE: PRT
;   ORGANISM: Aspergillus fumigatus
US-10-128-714-8252

  Query Match             28.9%;  Score 401;  DB 15;  Length 284;
  Best Local Similarity   37.3%;  Pred. No. 4.5e-33;
  Matches  101;  Conservative   45;  Mismatches  115;  Indels   10;  Gaps    4;

Qy         10 RVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTHSNQEVLQSC 69
              :: |:| | || ||  ||: :  :   :|  | | | |      || :|| || |   :
Db          8 KLAFIGGGNMASAIIGGLV-SQDMNPANITVSEPWDVNREKIAKLGVQTTTSNGEAAANA 66

Qy         70 LLVIFATKPHVLPAVLAEVAPV---VTTEHILVSVAAGVSLSTLEELLPPN----TRVLR 122
               :|| | ||     |  |:|      |:  ::||:|||::|::::| |  |      ::|
Db         67 DIVIIAVKPQTTKNVCQELATAWSQRTSLPVVVSIAAGITLNSMKEWLRTNDGRTAHIVR 126

Qy        123 VLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEE--VPEAYVDIHTGLSGSG 180
              |:|| | :|:|||  :     | : | :|:  ||::  :  |     |  :|: |||||||
Db        127 VMPNTPALVKEGASGLLASDDVTAEEKELIGALLQSVSKATEWVEKEELLDVVTGLSGSG 186

Qy        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240
               |:  |  | |   | :|:    | |:|||| ||  |||:   : ||||| :| :| ||
Db        187 PAYFFAMVEHLVASATALGLSKEQATRLAAQTCLGAGKMLVESSEEPAQLRKNVTSPNGT 246

Qy        241 TIYGLHALEQGGLRAATMSAVEAATCRAKEL 271
              |   |   |    :     ||:||| || ||
Db        247 THAALQTFESLNFKEIVDKAVQAATSRAAEL 277


RESULT 12
US-09-712-363-174
; Sequence 174, Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
;  APPLICANT: Eisenberg, David
;  APPLICANT:  Rotstein, Sergio H.
;  APPLICANT:  Marcotte, Edward M.
;  TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
;  TITLE OF INVENTION:  INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
;  FILE REFERENCE: 07419-032001
;  CURRENT APPLICATION NUMBER: US/09/712,363
;  CURRENT FILING DATE:  2000-11-13
;  PRIOR APPLICATION NUMBER: PCT/US00/02246
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: 60/179,531
;  PRIOR FILING DATE: 2000-02-01
;  PRIOR APPLICATION NUMBER: 60/117,844
;  PRIOR FILING DATE: 1999-01-29
```

```
;  PRIOR APPLICATION NUMBER: 60/118,206,
;  PRIOR FILING DATE: 1999-02-01
;  PRIOR APPLICATION NUMBER: 60/126,593
;  PRIOR FILING DATE: 1999-03-26
;  PRIOR APPLICATION NUMBER: 60/134,093
;  PRIOR FILING DATE: 1999-05-14
;  PRIOR APPLICATION NUMBER: 60/134,092
;  PRIOR FILING DATE: 1999-05-14
;  PRIOR APPLICATION NUMBER: 60/165,124
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/165,086
;  PRIOR FILING DATE: 1999-11-12
;  NUMBER OF SEQ ID NOS: 292
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 174
;   LENGTH: 295
;   TYPE: PRT
;   ORGANISM: Mycobacterium tuberculosis
US-09-712-363-174

  Query Match             26.4%;  Score 366;  DB 10;  Length 295;
  Best Local Similarity   31.6%;  Pred. No. 1.9e-29;
  Matches   90;  Conservative   55;  Mismatches  116;  Indels   24;  Gaps    3;

Qy         10 RVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTHSNQEVLQSC 69
              |:  :| | :  |:  ||:|||:     ::|    ||      |      | |  : :::
Db          7 RIAIIGGGSIGEALLSGLLRAGRQVKDLVVAERMPDRANYLAQTYSVLVT-SAADAVENA 65

Qy         70 LLVIFATKPHVLPAVLAEVAPVV------TTEHILVSVAAGVSLSTLEELLPPNTRVLRV 123
                |: | ||  :  |:|::|         : | : |:| ||::::  |  ||  | |:|
Db         66 TFVVVAVKPADVEPVIADLANATAAAENDSAEQVFVTVVAGITIAYFESKLPAGTPVVRA 125

Qy        124 LPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAF 183
              :||   :|  |   :|:|| |   : : :  | :| |    |||: :|  | :|||| |:
Db        126 MPNAAALVGAGVTALAKGRFVTPQQLEEVSALFDAVGGVLTVPESQLDAVTAVSGSGPAY 185

Qy        184 VCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQH---------------- 226
                  ||| :  | :|:   :|  :||||: |:| |||   :
Db        186 FFLLVEALVDAGVGVGLSRQVATDLAAQTMAGSAAMLLERMEQDQGGANGELMGLRVDLT 245

Qy        227 PAQLRSDVCTPGGTTIYGLHALEQGGLRAATMSAVEAATCRAKEL 271
               ::||: | :|||||   |  ||:|| | |  :||:||  |:::|
Db        246 ASRLRAAVTSPGGTTAAALRELERGGFRMAVDAAVQAAKSRSEQL 290


RESULT 13
US-09-738-626-3960
; Sequence 3960, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3960
;   LENGTH: 270
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-3960

  Query Match             26.2%;  Score 362.5;  DB 10;  Length 270;
  Best Local Similarity   30.6%;  Pred. No. 3.9e-29;
  Matches   83;  Conservative   63;  Mismatches  114;  Indels   11;  Gaps    4;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL----GCRTTHSNQEVL 66
              :  :| |::  |:  ||| |  :  |:|     |:|:    | |    |      | :
Db          4 IAVIGGGQIGEALVSGLI-AANMNPQNIRV---TNRSEERGQELRDRYGILNMTDNSQAA 59

Qy         67 QSCLLVIFATKPHVLPAVLAEVAPVV---TTEHILVSVAAGVSLSTLEELLPPNTRVLRV 123
                 :|     ||  :  ||:|:   :   : : ::||:|||:|:: :||       |:||
Db         60 DEADVVFLCVKPKFIVEVLSEITGTLDNNSAQSVVVSMAAGISIAAMEESASAGLPVVRV 119

Qy        124 LPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAF 183
              :|| | :| :|   : :||:| : : : :: ||   |   || |: :|  | :|||  |:
Db        120 MPNTPMLVGKGMSTVTKGRYVDAEQLEQVKDLLSTVGDVLEVAESDIDAVTAMSGSSPAY 179

Qy        184 VCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIY 243
              :   :||| |  | :|:| : | ::|  :  | | |:   |: |::||: | :| |||:
Db        180 LFLVTEALIEAGVNLGLPRATAKKLAVASFEGAATMMKETGKEPSELRAGVSSPAGTTVA 239

Qy        244 GLHALEQGGLRAATMSAVEAATCRAKELSRK 274
               :  ||: |:| |   | :|   |::|| ::
Db        240 AIRELEEESGIRGAFYRAAQACADRSEELGKR 270


RESULT 14
US-09-939-980-445
; Sequence 445, Application US/09939980
; Patent No. US20020082234A1
;  GENERAL INFORMATION:
;       APPLICANT: Black, Michael
;                  Burnham, Martin
;                  Hodgson, John
;                  Knowles, David
;                  Lonetto, Michael
;                  Nicholas, Richard
;                  Pratt, Julie
;                  Reichard, Richard
;                  Rosenberg, Martin
;                  Ward, Judith
;       TITLE OF INVENTION: No. US20020082234A1el Prokaryotic Polynucleotides,
;                           Polypeptides and Their Uses
;       NUMBER OF SEQUENCES: 534
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: SmithKline Beecham Corporation
;            STREET: 709 Swedeland Road
;            CITY: King of Prussia
;            STATE: PA
;            COUNTRY: USA
;            ZIP: 19406-0939
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/939,980
;            FILING DATE: 27-Aug-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/936,165
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Gimmi, Edward R
;            REGISTRATION NUMBER: 38,891
;            REFERENCE/DOCKET NUMBER: P50549
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 610-270-4478
```

```
;             TELEFAX: 610-270-5090
;             TELEX: <Unknown>
;    INFORMATION FOR SEQ ID NO: 445:
;         SEQUENCE CHARACTERISTICS:
;             LENGTH: 144 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;         MOLECULE TYPE: Protein
;         SEQUENCE DESCRIPTION: SEQ ID NO: 445:
US-09-939-980-445

  Query Match             8.4%;  Score 116;  DB 9;  Length 144;
  Best Local Similarity   26.8%;  Pred. No. 0.0004;
  Matches   30;  Conservative   24;  Mismatches   56;  Indels    2;  Gaps    2;

Qy         16 AGRMAGAIAQGLIRAGKVEAQHI-LASAPTDRNLCHF-QALGCRTTHSNQEVLQSCLLV1 73
              || || ||  |:| :  ::|  | | :   :: |  | : ||   :: :  :|:    |
Db         12 AGNMAQAIFTGIINSSNLDANDIYLTNKSNEQALKAFAEKLGVNYSYDDATLLKDADYVF 71

Qy         74 FATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLP 125
                ||||   |:   : | :| :   :|: ||:    : :.|       |::|
Db         72 LGTKPHDFDALATRIKPHITKDXCFMSIMAGIPTDYIXQQLECQNPXARIMP 123


RESULT 15
US-09-912-020-256
; Sequence 256, Application US/09912020
; Patent No. US20020045592A1
; GENERAL INFORMATION:
;  APPLICANT: Zyskind, Judith
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Trawick, John
;  APPLICANT:  Forsyth, R. Allyn
;  APPLICANT:  Froelich, Jamie M.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
;  TITLE OF INVENTION:  ESCHERICHIA COLI
;  FILE REFERENCE: ELITRA.001DV1
;  CURRENT APPLICATION NUMBER: US/09/912,020
;  CURRENT FILING DATE:  2001-07-23
;  PRIOR APPLICATION NUMBER: 09/492,709
;  PRIOR FILING DATE: 2000-01-27
;  PRIOR APPLICATION NUMBER: 60/117,405
;  PRIOR FILING DATE: 1999-01-27
;  NUMBER OF SEQ ID NOS: 485
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 256
;   LENGTH: 299
;   TYPE: PRT
;   ORGANISM: E. Coli
US-09-912-020-256

  Query Match             8.2%;  Score 114;  DB 9;  Length 299;
  Best Local Similarity   21.7%;  Pred. No. 0.0018;
  Matches   69;  Conservative   46;  Mismatches  119;  Indels   84;  Gaps   16;

Qy         10 RVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRN---LCHFQALGCRTTHSNQEVL 66
              :|||:| | |   ::: |::||       :    |||   :    | |  |  : : :
Db          7 KVGFIGLGIMGKPMSKNLLKAG-------YSLVVADRNPEAIADVIAAGAETASTAKAIA 59

Qy         67 QSC--LLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVL 124
              : |  :: :    |||    | |         |:     |  |  :  : |  :| :
Db         60 EQCDVIITMLPNSPHVKEVALGE-------NGIIEGAKPGTVLIDMSSIAPLASREI--- 109

Qy        125 PNLPCVVQEGAIVMARG-----RHVGSSETKLLQHLLEA-CGRCEEVPEAYVD------- 171
                       : |:|       |   | | :   |    |  : : : | |
Db        110 ---------SEALKAKGIDMLDAPVSGGEPKAIDGTLSVMVGGDKAIFDKYYDLMKAMAG 160

Qy        172 --IHTGLSGSG----------VAF-VCAFSEALAEGAVKMGMPSSLAHR-----IAAQTL 213
                :|||  |:|          ||  : | ||||   | | |:   | ::     :|  |:

Db        161 SVVHTGEIGAGNVTKLANQVIVALNIAAMSEALTL-ATKAGVNPDLVYQAIRGGLAGSTV 219

Qy        214 L-GTAKMLLHEGQHPA--------QLRSDVCTPGG---------TTIYGLHALEQGGLRA 255
              |   | |::     |          | : : |  |           :  : ||   ||
Db        220 LDAKAPMVMDRNFKPGFRIDLHIKDLANALDTSHGVGAQLPLTAAVMEMMQALRADGLGT 279

Qy        256 ATMSAVEAATCRAKELSR 273
              |  ||:    |  ::|::
Db        280 ADHSAL---ACYYEKLAK 294


Search completed: July 30, 2003, 15:29:11
Job time : 23 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 30, 2003, 15:23:05 ; Search time 85 Seconds
(without alignments)
511.660 Million cell updates/sec

Title: US-09-806-536A-14
Perfect score: 1386
Sequence: 1 MAAAEPSPRRVGFVGAGRMA.....AATMSAVEAATCRAKELSRK 274

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1386	100.0	274	21	AAY92517	Human OXRE-14. Ho
2	1386	100.0	274	22	AAB95591	Human protein sequ
3	1386	100.0	274	24	AAG79743	P5CR related polyp
4	1386	100.0	274	24	AAG79745	P5CR related polyp
5	1383	99.8	274	22	AAG66956	Human dihydropyrro
6	590.5	42.6	343	23	ABP41278	Human ovarian anti
7	584.5	42.2	319	22	AAB74780	Human P5CR protein
8	584.5	42.2	319	24	AAG79742	P5CR related polyp
9	584.5	42.2	319	24	AAG79744	P5CR related polyp
10	577.5	41.7	320	22	AAE12784	Human delta 1-pyrr
11	560.5	40.4	273	22	ABB61855	Drosophila melanog
12	560.5	40.4	315	22	AAE12785	Human delta 1-pyrr
13	560.5	40.4	320	22	AAB74779	Human Py-CR protei
14	537.5	38.8	314	21	AAB20584	Human delta 1-pyrr
15	537.5	38.8	314	23	AAU99322	Human delta 1-pyrr
16	505.5	36.5	234	24	ABR41145	Human DITHP enzyme
17	505.5	36.5	280	22	ABB71454	Drosophila melanog
18	477	34.4	319	24	AAG79746	P5CR related polyp
19	475	34.3	276	21	AAG06978	Arabidopsis thalia
20	469	33.8	266	23	ABB47427	Listeria monocytog
21	467.5	33.7	255	21	AAB56640	Human prostate can
22	436	31.5	256	21	AAG06979	Arabidopsis thalia
23	423.5	30.6	269	22	AAG64106	Escherichia coli p
24	401	28.9	284	24	ABJ25594	Aspergillus fumiga
25	401	28.9	284	24	ABJ26194	Aspergillus fumiga
26	395	28.5	265	24	ABU01336	S. pneumoniae type
27	392.5	28.3	268	22	AAU50524	Propionibacterium
28	388	28.0	212	22	AAB93027	Human protein sequ
29	366	26.4	295	22	AAG81123	Mycobacterium tube
30	362.5	26.2	270	22	AAG90206	C glutamicum prote
31	362.5	26.2	270	22	AAB79782	Corynebacterium gl
32	360.5	26.0	274	23	ABP73860	Candida albicans e
33	350.5	25.3	326	23	ABP65790	Bifidobacterium lo
34	349	25.2	262	23	ABB55233	Lactococcus lactis
35	344.5	24.9	256	23	ABP25621	Streptococcus poly
36	343	24.7	90	22	AAU23472	Novel human enzyme
37	325	23.4	83	22	AAU23493	Novel human enzyme
38	316.5	22.8	256	23	ABP25622	Streptococcus poly
39	309	22.3	151	21	AAG06980	Arabidopsis thalia
40	269	19.4	271	22	AAG82031	S. epidermidis ope
41	269	19.4	282	23	ABP38310	Staphylococcus epi
42	251.5	18.1	263	23	AAG66078	N. meningitidis pr
43	245.5	17.7	108	22	ABG20692	Novel human diagno
44	244.5	17.6	253	24	ABP80908	N. gonorrhoeae ami
45	202.5	14.6	279	23	ABB49027	Listeria monocytog

ALIGNMENTS

RESULT 1
AAY92517
ID AAY92517 standard; Protein; 274 AA.
XX
AC AAY92517;
XX
DT 10-AUG-2000 (first entry)
XX
DE Human OXRE-14.
XX
KW OXRE-14; oxidoreductase; pyrroline-5-carboxylate-reductase;
KW antiproliferative; anticancer; hepatotropic; antiviral; antiasthmatic;
KW nootropic; neuroprotectant; antiparkinsonian's; antisclerotic;
KW anxiolytic; antischizophrenic; anti-inflammatory; antiarthritic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 9..256
FT /label= signature_sequence
FT /note= "for delta 1-pyrroline-5-carboxylate reductase"
FT Modified-site 7
FT /note= "potential phosphorylation site"
FT Modified-site 44
FT /note= "potential phosphorylation site"
FT Modified-site 61
FT /note= "potential phosphorylation site"
FT Modified-site 108
FT /note= "potential phosphorylation site"
FT Modified-site 109
FT /note= "potential phosphorylation site"
FT Modified-site 216

```
FT                   /note= "potential phosphorylation site"
FT   Modified-site   259
FT                   /note= "potential phosphorylation site"
FT   Modified-site   265
FT                   /note= "potential phosphorylation site"
XX
PN   WO200020604-A2.
XX
PD   13-APR-2000.
XX
PF   06-OCT-1999;   99WO-US23434.
XX
PR   06-OCT-1998;   98US-0172227.
PR   02-DEC-1998;   98US-0155202.
PR   10-MAR-1999;   99US-0123911.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Lal P,  Guegler KJ,  Gorgone GA,  Corley NC,  Baughn MR,  Tang YT;
PI   Hillman JL,  Bandman O,  Azimzai Y,  Au-Young J,  Yue H,  Lu DAM;
PI   Yang J;
XX
DR   WPI; 2000-303785/26.
DR   N-PSDB; AAA09388.
XX
PT   Purified polypeptide for treating or preventing disorders associated
PT   with decreased expression or activity of oxidoreductase molecules
XX
PS   Claim 1; Page 84; 97pp; English.
XX
CC   This OXRE-14 has identity with pyrroline-5-carboxylate-reductase.
CC   The polypeptides are useful for treating or preventing a disorder
CC   associated with decreased expression or activity of OXRE.  Antagonists of
CC   OXRE are useful for treating or preventing a disorder associated with
CC   increased expression or activity of OXRE.  The disorders include cell
CC   proliferative disorders (cirrhosis, hepatitis), cancer (leukemia,
CC   melanoma), hypopituitarism and hyperpituitarism, hypothyroidism and
CC   hyperthyroidism, metabolic disorders (Addison's disease, cystic
CC   fibrosis), reproductive disorders (infertility, ovulatory defects),
CC   neurological disorders (Alzheimer's disease, Parkinson's disease,
CC   multiple sclerosis), mental disorders (anxiety, schizophrenia),
CC   autoimmune/inflammatory disorders (acquired immunodeficiency syndrome
CC   (AIDS), asthma, osteoarthritis), and viral infections.  The
CC   polynucleotides may be used in Southern or Northern analysis, polymerase
CC   chain reaction (PCR), or in enzyme-linked immunosorbent assays (ELISA).
XX
SQ   Sequence   274 AA;

  Query Match             100.0%;  Score 1386;  DB 21;  Length 274;
  Best Local Similarity   100.0%;  Pred. No. 4.7e-131;
  Matches  274;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60

Qy         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120

Qy        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180

Qy        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240

Qy        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274
              ||||||||||||||||||||||||||||||||||
Db        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274
```

```
RESULT 2
AAB95591
ID   AAB95591 standard; Protein; 274 AA.
XX
AC   AAB95591;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human protein sequence SEQ ID NO:18269.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 8; SEQ ID 18269; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence   274 AA;

  Query Match             100.0%;  Score 1386;  DB 22;  Length 274;
  Best Local Similarity   100.0%;  Pred. No. 4.7e-131;
  Matches  274;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60
```

```
Qy         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120

Qy        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180

Qy        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240

Qy        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274
              ||||||||||||||||||||||||||||||||||
Db        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274
```

RESULT 3
AAG79743
ID AAG79743 standard; Protein; 274 AA.
XX
AC AAG79743;
XX
DT 18-MAR-2003 (first entry)
XX
DE P5CR related polypeptide #2.
XX
KW Pyrroline 5 carboxylate reductase; P5CR; NADPH; cancer; breast; colon;
KW pyrroline 5 carboxylate; P5C; proline; p53; kidney; lung; ovary.
XX
OS Homo sapiens.
XX
PN WO200299043-A2.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-US17319.
XX
PR 05-JUN-2001; 2001US-296080P.
PR 10-OCT-2001; 2001US-328509P.
XX
PA (EXEL-) EXELIS INC.
XX
PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
PI Engst S;
XX
DR WPI; 2003-140606/13.
XX
PT Identifying p53 pathway modulating agents with P5CR genes, useful for
PT the diagnosis and treatment of disorders associated with defects in the
PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and
PT ovary -
XX
PS Claim 13; Page 48-49; 53pp; English.
XX
CC The sequences given in AAG79742-46 represent pyrroline 5 carboxylate
CC reductase (P5CR) related polypeptide sequences. P5CR catalyses the
CC NAD(P)H-dependent conversion of pyrroline 5 carboxylate (P5C) to
CC proline. These sequences may be used in the method of the invention
CC for identifying a candidate p53 pathway modulating agent. The method
CC comprises providing an assay system comprising a purified P5CR
CC polypeptide or nucleic acid, or a functionally active fragment or
CC derivative, contacting the assay system with a test agent, where the
CC system provides a reference activity, and detecting a test agent-biased
CC activity of the assay system. The method of the present invention is
CC useful for the diagnosis and treatment of disorders associated with
CC defects in the p53 pathway, such as cancer of the breast, colon,
CC kidneys, lung and ovary.
XX
SQ Sequence 274 AA;

Query Match 100.0%; Score 1386; DB 24; Length 274;
Best Local Similarity 100.0%; Pred. No. 4.7e-131;

Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60

Qy         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120

Qy        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180

Qy        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240

Qy        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274
              ||||||||||||||||||||||||||||||||||
Db        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274
```

RESULT 4
AAG79745
ID AAG79745 standard; Protein; 274 AA.
XX
AC AAG79745;
XX
DT 18-MAR-2003 (first entry)
XX
DE P5CR related polypeptide #4.
XX
KW Pyrroline 5 carboxylate reductase; P5CR; NADPH; cancer; breast; colon;
KW pyrroline 5 carboxylate; P5C; proline; p53; kidney; lung; ovary.
XX
OS Homo sapiens.
XX
PN WO200299043-A2.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-US17319.
XX
PR 05-JUN-2001; 2001US-296080P.
PR 10-OCT-2001; 2001US-328509P.
XX
PA (EXEL-) EXELIS INC.
XX
PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
PI Engst S;
XX
DR WPI; 2003-140606/13.
XX
PT Identifying p53 pathway modulating agents with P5CR genes, useful for
PT the diagnosis and treatment of disorders associated with defects in the
PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and
PT ovary -
XX
PS Claim 13; Page 50-52; 53pp; English.
XX
CC The sequences given in AAG79742-46 represent pyrroline 5 carboxylate
CC reductase (P5CR) related polypeptide sequences. P5CR catalyses the
CC NAD(P)H-dependent conversion of pyrroline 5 carboxylate (P5C) to
CC proline. These sequences may be used in the method of the invention
CC for identifying a candidate p53 pathway modulating agent. The method
CC comprises providing an assay system comprising a purified P5CR
CC polypeptide or nucleic acid, or a functionally active fragment or
CC derivative, contacting the assay system with a test agent, where the
CC system provides a reference activity, and detecting a test agent-biased
CC activity of the assay system. The method of the present invention is
CC useful for the diagnosis and treatment of disorders associated with
CC defects in the p53 pathway, such as cancer of the breast, colon,

```
CC   kidneys, lung and ovary.
XX
SQ   Sequence   274 AA;

  Query Match           100.0%;  Score 1386;  DB 24;  Length 274;
  Best Local Similarity  100.0%;  Pred. No. 4.7e-131;
  Matches  274;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60

Qy         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120

Qy        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180

Qy        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240

Qy        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274
              ||||||||||||||||||||||||||||||||||
Db        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274


RESULT 5
AAG66956
ID   AAG66956 standard; Protein; 274 AA.
XX
AC   AAG66956;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human dihydropyrrole-5-carboxylate reductase 30 polypeptide.
XX
KW   Human; dihydropyrrole-5-carboxylate reductase 30; cancer; cytostatic;
KW   human immunodeficiency virus; HIV; infection; immunological disease;
KW   inflammatory disease.
XX
OS   Homo sapiens.
XX
PN   CN1298002-A.
XX
PD   06-JUN-2001.
XX
PF   24-NOV-1999;   99CN-0124090.
XX
PR   24-NOV-1999;   99CN-0124090.
XX
PA   (SHAN-) SHANGHAI BORONG GENE DEV CO LTD.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2001-489680/54.
DR   N-PSDB; AAH77597.
XX
PT   Human dihydropyrrole-5-carboxylate reductase 30 as one new kind of
PT   polypeptide and polynucleotides encoding this polypeptide -
XX
PS   Claim 1; Page 20-21 (disclosure); 26pp; Chinese.
XX
CC   The invention relates to a novel polypeptide, human
CC   dihydropyrrole-5-carboxylate reductase 30, polynucleotides encoding
CC   this polypeptide and a DNA recombination process to produce the
CC   polypeptide. The polypeptide is useful for treating various diseases,
CC   such as malignant tumours, nosohaemia, HIV infection, immunological
CC   diseases and inflammatory diseases. The invention also provides an
CC   antibody against the polypeptide. The present sequence is the
CC   polypeptide of the invention.
XX
SQ   Sequence   274 AA;

  Query Match           99.8%;  Score 1383;  DB 22;  Length 274;
  Best Local Similarity  99.6%;  Pred. No. 9.4e-131;
  Matches  273;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60

Qy         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGMSLSTLEELLPPNTRV 120

Qy        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180

Qy        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240

Qy        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274
              ||||||||||||||||||||||||||||||||||
Db        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274


RESULT 6
ABP41278
ID   ABP41278 standard; Protein; 343 AA.
XX
AC   ABP41278;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Human ovarian antigen HOGDC67, SEQ ID NO:2410.
XX
KW   Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW   ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW   infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW   PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW   inflammatory condition; immune disorder; blood disorder;
KW   cardiovascular disorder; respiratory disorder; neurological disorder;
KW   gastrointestinal disorder; urinary system disorder; drug screening;
KW   gene therapy; chromosome mapping; forensic analysis;
KW   antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW   antiinflammatory; gynaecological; reproductive; chromosome 17.
XX
OS   Homo sapiens.
XX
PN   WO200200677-A1.
XX
PD   03-JAN-2002.
XX
PF   07-JUN-2001; 2001WO-US18569.
XX
PR   07-JUN-2000; 2000US-209467P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Birse CE,  Rosen CA;
XX
DR   WPI; 2002-147878/19.
DR   N-PSDB; ABQ54355.
XX
PT   Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT   useful in the prevention, treatment and diagnosis of cancer (e.g.
PT   ovarian cancer), immune disorders, cardiovascular disorders and
PT   neurological diseases -
XX
PS   Claim 11; SEQ ID No 2410; 2922pp; English.
XX
```

CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents a human ovarian antigen of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 343 AA;

Query Match 42.6%; Score 590.5; DB 23; Length 343;
Best Local Similarity 45.1%; Pred. No. 8.9e-51;
Matches 125; Conservative 49; Mismatches 96; Indels 7; Gaps 3;

```
Qy          1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCR 57
              :    |    |||:|||::| |:|:|   || : |  |:||:| | :|    ||   | :
Db         17 LGQGSPDSMSVGFIGAGQLAFALAKGFTAAGVLAAHKIMASSP-DMDLATVSALRKMGVK 75

Qy         58 TTHSNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLP-- 115
                |  |:| :|   ::  | |||::| :| |:   :   ||:|| ||||::|::|: |
Db         76 LTPHNKETVQHSDVLFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAF 135

Qy        116 -PNTRVLRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHT 174
               |  ||:| : | | ||:||| | | |    : :|:: || : | | || |  :|  |
Db        136 RPAPRVIRCMTNTPVVVREGATVYATGTHAQVEDGRLMEQLLSSVGFCTEVEEDLIDAVT 195

Qy        175 GLSGSGVAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDV 234
              |||||| |:     :|||:| ||||:|  || |: || ||| ||||||  ||| ||: :|
Db        196 GLSGSGPAYAFTALDALADGGVKMGLPRRLAVRLGAQALLGAAKMLLHSEQHPGQLKDNV 255

Qy        235 CTPGGTTIYGLHALEQGGLRAATMSAVEAATCRAKEL 271
               :||| ||: || || || |:  ::||||:  | :||
Db        256 SSPGGATIHALHVLESGGFRSLLINAVEASCIRTREL 292
```

RESULT 7
AAB74780
ID AAB74780 standard; Protein; 319 AA.
XX
AC AAB74780;
XX
DT 06-JUN-2001 (first entry)
XX
DE Human P5CR protein.
XX
KW Human; pyrroline 5'-carboxylate reductase; Py-CR; P5CR.
XX
OS Homo sapiens.
XX
PN CN1274728-A.
XX
PD 29-NOV-2000.
XX
PF 25-MAY-1999; 99CN-0107071.
XX
PR 25-MAY-1999; 99CN-0107071.
XX
PA (UYFU-) UNIV FUDAN.
XX
PI Yu L, Fu Q, Zhang H;
XX
DR WPI; 2001-211749/22.
XX
PT New human protein and its code sequence, preparation and application -
XX
PS Example 2; Fig 1; 20pp; Chinese.
XX
CC The present invention describes a human protein designated Py-CR, which
CC is homologous to the human pyrroline-5'-carboxylate reductase (P5CR)
CC (EC1.5 1.2). The present invention also describes methods for the
CC application and production process of the Py-CR polynucleotide and
CC protein sequences. The present sequence represents the human P5CR
CC protein which is used in comparison with the human Py-CR protein in
CC an example from the present invention.
XX
SQ Sequence 319 AA;

Query Match 42.2%; Score 584.5; DB 22; Length 319;
Best Local Similarity 46.4%; Pred. No. 3.2e-50;
Matches 124; Conservative 47; Mismatches 89; Indels 7; Gaps 3;

```
Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
              |||:|||::| |:|:|   || : |  |:||:| | :|    ||   | : |  |:| :|
Db          3 VGFIGAGQLAFALAKGFTAAGVLAAHKIMASSP-DMDLATVSALRKMGVKLTPHNKETVQ 61

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLP---PNTRVLRVL 124
                ::  | |||::| :| |:   :   ||:|| ||||::|::|: |    |  ||:| :
Db         62 HSDVLFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRPAPRVIRCM 121

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
               | | ||:||| | | | |    : :|:: ||   | | || |  :|  |||||| |:
Db        122 TNTPVVVREGATVYATGTHAQVEDGRLMEQLLSTVGFCTEVEEDLIDAVTGLSGSGPAYA 181

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
                :|||:| ||||:|  || |: || ||| ||||||  ||| ||: :| :||| ||:
Db        182 FTALDALADGGVKMGLPRRLAVRLGAQALLGAAKMLLHSEQHPGQLKDNVSSPGGATIHA 241

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
              || || || |:  ::||||:  | :||
Db        242 LHVLESGGFRSLLINAVEASCIRTREL 268
```

RESULT 8
AAG79742
ID AAG79742 standard; Protein; 319 AA.
XX
AC AAG79742;
XX
DT 18-MAR-2003 (first entry)
XX
DE P5CR related polypeptide #1.
XX
KW Pyrroline 5 carboxylate reductase; P5CR; NADPH; cancer; breast; colon;
KW pyrroline 5 carboxylate; P5C; proline; p53; kidney; lung; ovary.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 1..253
FT /label= P5CR domain
XX
PN WO200299043-A2.
XX
PD 12-DEC-2002.

XX
PF 03-JUN-2002; 2002WO-US17319.
XX
PR 05-JUN-2001; 2001US-296080P.
PR 10-OCT-2001; 2001US-328509P.
XX
PA (EXEL-) EXELIS INC.
XX
PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
PI Engst S;
XX
DR WPI; 2003-140606/13.
XX
PT Identifying p53 pathway modulating agents with P5CR genes, useful for
PT the diagnosis and treatment of disorders associated with defects in the
PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and
PT ovary -
XX
PS Claim 13; Page 46-48; 53pp; English.
XX
CC The sequences given in AAG79742-46 represent pyrroline 5 carboxylate
CC reductase (P5CR) related polypeptide sequences. P5CR catalyses the
CC NAD(P)H-dependent conversion of pyrroline 5 carboxylate (P5C) to
CC proline. These sequences may be used in the method of the invention
CC for identifying a candidate p53 pathway modulating agent. The method
CC comprises providing an assay system comprising a purified P5CR
CC polypeptide or nucleic acid, or a functionally active fragment or
CC derivative, contacting the assay system with a test agent, where the
CC system provides a reference activity, and detecting a test agent-biased
CC activity of the assay system. The method of the present invention is
CC useful for the diagnosis and treatment of disorders associated with
CC defects in the p53 pathway, such as cancer of the breast, colon,
CC kidneys, lung and ovary.
XX
SQ Sequence 319 AA;

Query Match 42.2%; Score 584.5; DB 24; Length 319;
Best Local Similarity 46.4%; Pred. No. 3.2e-50;
Matches 124; Conservative 47; Mismatches 89; Indels 7; Gaps 3;

```
Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
              |||:|||::| |:|:|   || : |  |:||:| | :|    ||   | : |  |:| :|
Db          3 VGFIGAGQLAFALAKGFTAAGVLAAHKIMASSP-DMDLATVSALRKMGVKLTPHNKETVQ 61

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLP---PNTRVLRVL 124
                ::  | |||::| :| |:   :   ||:|| ||||::|::|: |    |  ||:| :
Db         62 HSDVLFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRPAPRVIRCM 121

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
              | | ||:||| | | | |    : :|:: ||   | | || |  :|  |||||| |:
Db        122 TNTPVVVREGATVYATGTHAQVEDGRLMEQLLSTVGFCTEVEEDLIDAVTGLSGSGPAYA 181

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
                :|||:| ||||:|  || |: || ||| ||||||  ||| ||: :| :||| ||:
Db        182 FTALDALADGGVKMGLPRRLAVRLGAQALLGAAKMLLHSEQHPGQLKDNVSSPGGATIHA 241

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
              || || || |:  ::|||:  | :||
Db        242 LHVLESGGFRSLLINAVEASCIRTREL 268
```

RESULT 9
AAG79744
ID AAG79744 standard; Protein; 319 AA.
XX
AC AAG79744;
XX
DT 18-MAR-2003 (first entry)
XX
DE P5CR related polypeptide #3.
XX
KW Pyrroline 5 carboxylate reductase; P5CR; NADPH; cancer; breast; colon;
KW pyrroline 5 carboxylate; P5C; proline; p53; kidney; lung; ovary.
XX
OS Homo sapiens.
XX
PN WO200299043-A2.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-US17319.
XX
PR 05-JUN-2001; 2001US-296080P.
PR 10-OCT-2001; 2001US-328509P.
XX
PA (EXEL-) EXELIS INC.
XX
PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
PI Engst S;
XX
DR WPI; 2003-140606/13.
XX
PT Identifying p53 pathway modulating agents with P5CR genes, useful for
PT the diagnosis and treatment of disorders associated with defects in the
PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and
PT ovary -
XX
PS Claim 13; Page 49-50; 53pp; English.
XX
CC The sequences given in AAG79742-46 represent pyrroline 5 carboxylate
CC reductase (P5CR) related polypeptide sequences. P5CR catalyses the
CC NAD(P)H-dependent conversion of pyrroline 5 carboxylate (P5C) to
CC proline. These sequences may be used in the method of the invention
CC for identifying a candidate p53 pathway modulating agent. The method
CC comprises providing an assay system comprising a purified P5CR
CC polypeptide or nucleic acid, or a functionally active fragment or
CC derivative, contacting the assay system with a test agent, where the
CC system provides a reference activity, and detecting a test agent-biased
CC activity of the assay system. The method of the present invention is
CC useful for the diagnosis and treatment of disorders associated with
CC defects in the p53 pathway, such as cancer of the breast, colon,
CC kidneys, lung and ovary.
XX
SQ Sequence 319 AA;

Query Match 42.2%; Score 584.5; DB 24; Length 319;
Best Local Similarity 46.4%; Pred. No. 3.2e-50;
Matches 124; Conservative 47; Mismatches 89; Indels 7; Gaps 3;

```
Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
              |||:|||::| |:|:|   || : |  |:|||:| | :|    ||   | : |  |:| :|
Db          3 VGFIGAGQLAFALAKGFTAAGVLAAHKIMASSP-DMDLATVSALRKMGVKLTPHNKETVQ 61

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLP---PNTRVLRVL 124
                ::  | |||::| :| |:   :   ||:|| ||||::|::|: |    |  ||:| :
Db         62 HSDVLFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRPAPRVIRCM 121

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
              | | ||:||| | | | |    : :|:: ||   | | || |  :|  |||||| |:
Db        122 TNTPVVVREGATVYATGTHAQVEDGRLMEQLLSTVGFCTEVEEDLIDAVTGLSGSGPAYA 181

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
                :|||:| ||||:|  || |: || ||| ||||||  ||| ||: :| :||| ||:
Db        182 FTALDALADGGVKMGLPRRLAVRLGAQALLGAAKMLLHSEQHPGQLKDNVSSPGGATIHA 241

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
              || || || |:  ::|||:  | :||
Db        242 LHVLESGGFRSLLINAVEASCIRTREL 268
```

RESULT 10
AAE12784
ID AAE12784 standard; Protein; 320 AA.
XX
AC AAE12784;
XX

```
DT   15-JAN-2002  (first entry)
XX
DE   Human delta 1-pyrroline-5-carboxylate reductase homologue (P5CRH).
XX
KW   Human; delta 1-pyrroline-5-carboxylate reductase; P5CR; cystic fibrosis;
KW   osteoporosis; neuronal disorder; gene therapy; akathesia; drug screening;
KW   actinic keratosis; Alzheimer's disease; amyotrophic lateral sclerosis;
KW   connective tissue disorder; myocardial fibrosis; cell proliferation;
KW   arteriosclerosis bursitis; cancer; amnesia; neuroprotective; cytostatic;
KW   antiarteriosclerotic; osteopathic; cardiant.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..18
FT                   /label= Signal_peptide
FT   Protein         19..320
FT                   /label= Mature_human P5CRH
FT   Misc-difference 14
FT                   /note= "This amino acid residue is absent in the
FT                   sequence shown in sequence listing of the specification"
FT   Misc-difference 15
FT                   /note= "This amino acid residue is absent in the
FT                   sequence shown in sequence listing of the specification"
FT   Misc-difference 16
FT                   /note= "This amino acid residue is absent in the
FT                   sequence shown in sequence listing of the specification"
FT   Misc-difference 18
FT                   /note= "This amino acid residue is absent in the
FT                   sequence shown in sequence listing of the specification"
FT   Misc-difference 188
FT                   /note= "This amino acid residue is absent in the
FT                   sequence shown in sequence listing of the specification"
FT   Misc-difference 189
FT                   /note= "This amino acid residue is absent in the
FT                   sequence shown in sequence listing of the specification"
XX
PN   US6268192-B1.
XX
PD   31-JUL-2001.
XX
PF   05-MAY-2000; 2000US-0565910.
XX
PR   18-JUN-1998;   98US-0099676.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Hillman JL,  Corley NC,  Baughn MR;
XX
DR   WPI; 2001-647056/74.
DR   N-PSDB; AAD20894.
XX
PT   New delta 1-pyrroline-5-carboxylate reductase polypeptides and
PT   polynucleotides, useful for diagnosing, treating and preventing
PT   neuronal disorders, connective tissue disorders or disorders of cell
PT   proliferation  -
XX
PS   Claim 1; Fig 2; 32pp; English.
XX
CC   The present invention relates to delta 1-pyrroline-5-carboxylate
CC   reductase (P5CR) polypeptides and polynucleotides. The P5CR gene is used
CC   in gene therapy. The P5CR nucleic acid and amino acid sequences are
CC   useful in the diagnosis, treatment and prevention of neuronal disorders
CC   (e.g. akathesia, Alzheimer's disease, amnesia, or amyotrophic lateral
CC   sclerosis), connective tissue disorders (e.g. cystic fibrosis, myocardial
CC   fibrosis, or osteoporosis) and disorders of cell proliferation (e.g.
CC   actinic keratosis, arteriosclerosis bursitis, cancers). P5CR may also be
CC   used for screening libraries of compounds in various drug screening
CC   techniques. The present sequence is human delta 1-pyrroline-5-carboxylate
CC   reductase homologue (P5CRH).
XX
SQ   Sequence   320 AA;
```

```
  Query Match             41.7%;  Score 577.5;  DB 22;  Length 320;
  Best Local Similarity   46.4%;  Pred. No. 1.6e-49;
  Matches  124;  Conservative   48;  Mismatches   88;  Indels    7;  Gaps    3;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
              |||:|||::| |:|:|   || : |  |:||:| : ||    ||   |   | ||:| ::
Db          3 VGFIGAGQLAYALARGFTAAGILSAHKIIASSP-EMNLPTVSALRKMGVNLTRSNKETVK 61

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELL---PPNTRVLRVL 124
                ::  | |||::| :| |:   |   ||:|| ||||::|::|: |    | :|:| :
Db         62 HSDVLFLAVKPHIIPFILDEIGADVQARHIVVSCAAGVTISSVEKKLMAFQPAPKVIRCM 121

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
               | | |||||| | | | |    : :||: |: : | | || |  :|  ||||||| |:
Db        122 TNTPVVVQEGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGSGPAYA 181

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
                :|||:| ||||:|  || :: || ||| |||||   ||| ||: :||:||| ||:
Db        182 FMALDALADGGVKMGLPRRLAIQLGAQALLGAAKMLLDSEQHPCQLKDNVCSPGGATIHA 241

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
              || || || |:  ::|||||:  | :||
Db        242 LHFLESGGFRSLLINAVEASCIRTREL 268
```

```
RESULT 11
ABB61855
ID   ABB61855 standard; Protein; 273 AA.
XX
AC   ABB61855;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 12357.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL05958.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 12357; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
```

```
XX
SQ   Sequence   273 AA;

  Query Match          40.4%;  Score 560.5;  DB 22;  Length 273;
  Best Local Similarity   44.2%;  Pred. No. 6.8e-48;
  Matches  118;  Conservative   50;  Mismatches   96;  Indels    3;  Gaps    3;

Qy         10 RVGFVGAGRMAGAIAQGLIRAGKVEAQHILASA-PTDR-NLCHFQALGCRTTHSNQEVLQ 67
              ::||:| | || |:|:| : ||  :   ::||  | |: :|  ||:|| |   |  |:|
Db          6 KIGFLGGGNMAKALAKGFLAAGLAKPNTLIASVHPADKLSLQSFQSLGVETVIKNAPVVQ 65

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNL 127
                :|  : || |:|:|||:|: | :::  : :||| |::|||:|  | |  ||:|||:|||
Db         66 QSDVVFVSVKPQVVPSVLSEIQP-LSSGKLFLSVAMGITLSTIESSLSPQARVIRVMPNL 124

Qy        128 PCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAF 187
              | ||  |  |  ||      ::  : | ||:: | || | |: :|: | |||| |:|
Db        125 PAVVCSGCSVFVRGSKATDADADITQKLLQSVGTCEPVDESQLDVVTALSGSGPAYVFVM 184

Qy        188 SEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHA 247
               ||||:||| ||||  ||:|:|:||:||   |:   | || ||:  | :| |:|   |
Db        185 IEALADGAVHMGMPRDLAYRLASQTVLGAGHMVRDSGMHPGQLKDGVTSPAGSTAAALRQ 244

Qy        248 LEQGGLRAATMSAVEAATCRAKELSRK 274
              ||  | |||   ||| || | :::| |
Db        245 LELSGFRAAVSGAVEQATLRCRQISGK 271


RESULT 12
AAE12785
ID   AAE12785 standard; Protein; 315 AA.
XX
AC   AAE12785;
XX
DT   15-JAN-2002  (first entry)
XX
DE   Human delta 1-pyrroline-5-carboxylate reductase (P5CR).
XX
KW   Human; delta 1-pyrroline-5-carboxylate reductase; P5CR; cystic fibrosis;
KW   osteoporosis; neuronal disorder; gene therapy; akathesia; drug screening;
KW   actinic keratosis; Alzheimer's disease; amyotrophic lateral sclerosis;
KW   connective tissue disorder; myocardial fibrosis; cell proliferation;
KW   arteriosclerosis bursitis; cancer; amnesia; neuroprotective; cytostatic;
KW   antiarteriosclerotic; osteopathic; cardiant.
XX
OS   Homo sapiens.
XX
PN   US6268192-B1.
XX
PD   31-JUL-2001.
XX
PF   05-MAY-2000; 2000US-0565910.
XX
PR   18-JUN-1998;   98US-0099676.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Hillman JL,  Corley NC,  Baughn MR;
XX
DR   WPI; 2001-647056/74.
XX
PT   New delta 1-pyrroline-5-carboxylate reductase polypeptides and
PT   polynucleotides, useful for diagnosing, treating and preventing
PT   neuronal disorders, connective tissue disorders or disorders of cell
PT   proliferation  -
XX
PS   Disclosure; Fig 2; 32pp; English.
XX
CC   The present invention relates to delta 1-pyrroline-5-carboxylate
CC   reductase (P5CR) polypeptides and polynucleotides. The P5CR gene is used
CC   in gene therapy. The P5CR nucleic acid and amino acid sequences are
CC   useful in the diagnosis, treatment and prevention of neuronal disorders
CC   (e.g. akathesia, Alzheimer's disease, amnesia, or amyotrophic lateral
CC   sclerosis), connective tissue disorders (e.g. cystic fibrosis, myocardial
CC   fibrosis, or osteoporosis) and disorders of cell proliferation (e.g.
CC   actinic keratosis, arteriosclerosis bursitis, cancers). P5CR may also be
CC   used for screening libraries of compounds in various drug screening
CC   techniques. The present sequence is human delta 1-pyrroline-5-carboxylate
CC   reductase (P5CR).
XX
SQ   Sequence   315 AA;

  Query Match          40.4%;  Score 560.5;  DB 22;  Length 315;
  Best Local Similarity   46.1%;  Pred. No. 8.3e-48;
  Matches  123;  Conservative   46;  Mismatches   87;  Indels   11;  Gaps    5;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
              |||:|||::  | |:|   || : |  |:||:| | :|    ||   | : |  |:| :|
Db          3 VGFIGAGQL--AFAKGFTAAGVLAAHKIMASSP-DMDLATVSALRKMGVKLTPHNKETVQ 59

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLP---PNTRVLRVL 124
                 ::  | |||::| :| |:   :   ||:|| ||||::|:::|: |    |  ||:| :
Db         60 HSDVLFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRPAPRVIRCM 119

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
              | | ||:||| | | | |    : :|:: ||   | | || |  :|  |||||||  |:
Db        120 TNTPVVVREGATVYATGTHAQVEDGRLMEQLLSTVGFCTEVEEDLIDAVTGLSGSGPAY- 178

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
               ||:   |:| ||||:|  || |: || ||| ||||||  ||| ||: :| :||| ||:
Db        179 -AFTALDADGGVKMGLPRRLAVRLGAQALLGAAKMLLHSEQHPGQLKDNVSSPGGATIHA 237

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
              || || || |:  ::|||||:  | :||
Db        238 LHVLESGGFRSLLINAVEASCIRTREL 264


RESULT 13
AAB74779
ID   AAB74779 standard; Protein; 320 AA.
XX
AC   AAB74779;
XX
DT   06-JUN-2001  (first entry)
XX
DE   Human Py-CR protein SEQ ID NO:4.
XX
KW   Human; pyrroline 5'-carboxylate reductase; Py-CR; P5CR.
XX
OS   Homo sapiens.
XX
PN   CN1274728-A.
XX
PD   29-NOV-2000.
XX
PF   25-MAY-1999;   99CN-0107071.
XX
PR   25-MAY-1999;   99CN-0107071.
XX
PA   (UYFU-) UNIV FUDAN.
XX
PI   Yu L,  Fu Q,  Zhang H;
XX
DR   WPI; 2001-211749/22.
DR   N-PSDB; AAF81847.
XX
PT   New human protein and its code sequence, preparation and application -
XX
PS   Claim 4; Page 16; 20pp; Chinese.
XX
CC   The present invention describes a human protein designated Py-CR, which
CC   is homologous to the human pyrroline-5'-carboxylate reductase (P5CR)
CC   (EC1.5 1.2). The present invention also describes methods for the
CC   application and production process of the Py-CR polynucleotide and
CC   protein sequences. The present sequence represents the human Py-CR
CC   protein as given in the present invention.
```

```
XX
SQ   Sequence   320 AA;

  Query Match             40.4%;  Score 560.5;  DB 22;  Length 320;
  Best Local Similarity   45.7%;  Pred. No. 8.5e-48;
  Matches  122;  Conservative   48;  Mismatches   90;  Indels    7;  Gaps    3;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
              |||:|||::| |:|:|   || : |  |:||:| : ||    ||   |   | ||:| ::
Db          3 VGFIGAGQLANALARGFTAAGILSAHKIIASSP-EMNLPTVSALRKMGVNLTRSNKETVK 61

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELL---PPNTRVLRVL 124
                ::  | |  ::| :| |:   |   ||:|| ||||::|::|: |    |  :|:| :
Db         62 HSDVLFLAVKHIIIPFILDEIGADVQARHIVVSCAAGVTISSVEKKLMAFQPAPKVIRCM 121

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
               | | |||||| | | | |    : :||: |: : | | || |  :|  ||||||| |:
Db        122 TNTPVVVQEGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGSGPAYA 181

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
                :|||:| ||||:|  || :: || ||| |||||   ||| ||: :||:||| ||:
Db        182 FMALDALADGGVKMGLPRRLAIQLGAQALLGAAKMLLDSEQHPCQLKDNVCSPGGATIHA 241

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
              || || || |:  ::|||:  | :||
Db        242 LHFLESGGFRSLLINAVEASCIRTREL 268


RESULT 14
AAB20584
ID   AAB20584 standard; Protein; 314 AA.
XX
AC   AAB20584;
XX
DT   13-DEC-2000  (first entry)
XX
DE   Human delta 1-pyrroline-5-carboxylate reductase homologue protein.
XX
KW   Human; delta 1-pyrroline-5-carboxylate reductase homologue; P5CRH;
KW   nootropic; neuroprotective; osteopathic; cytostatic; antiParkinsonian;
KW   cardiant; osteopathic; ophthalmological; hepatotropic; gene therapy;
KW   neuronal disorder; connective tissue disorder; cell proliferation;
KW   akathesia; Alzheimer's disease; amnesia; dementia; Parkinson's disease;
KW   cystic fibrosis; myocardial fibrosis; osteoporosis; Marfan syndrome;
KW   cirrhosis, leukaemia, lymphoma, sarcoma, bone marrow cancer.
XX
OS   Synthetic.
XX
PN   US6100075-A.
XX
PD   08-AUG-2000.
XX
PF   18-JUN-1998;   98US-0099676.
XX
PR   18-JUN-1998;   98US-0099676.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Hillman JL,  Baughn MR,  Corley NC;
XX
DR   WPI; 2000-548292/50.
DR   N-PSDB; AAA88073.
XX
PT   New human delta 1-pyrroline-5-carboxylate reductase homolog (P5CRH) and
PT   polynucleotides encoding P5CRH, useful for diagnosing, treating or
PT   preventing neuronal or connective tissue disorders, or disorders of
PT   cell proliferation  -
XX
PS   Claim 1; Fig 1A-E; 32pp; English.
XX
CC   The present sequence represents human delta 1-pyrroline-5-carboxylate
CC   reductase homologue (P5CRH). P5CRH has nootropic, neuroprotective,
CC   osteopathic, cytostatic, antiParkinsonian, cardiant, osteopathic,
CC   ophthalmological and hepatotropic activities, and can be used in gene
CC   therapy. P5CRH and the polynucleotides encoding it are useful for
CC   diagnosing, treating or preventing neuronal disorders, connective tissue
CC   disorders, or disorders of cell proliferation. These may be applied to
CC   any subject, e.g. dogs, cats, cows, horses, rabbits, monkeys, or
CC   preferably humans. The polynucleotide sequences may be used to detect
CC   and quantitate gene expression in biopsied tissues where expression of
CC   the polypeptide may correlate with the disease. The diagnostic assay may
CC   be used to determine absence, presence, and excess expression of the
CC   polypeptide, and to monitor regulation of the polypeptide levels during
CC   therapeutic intervention. The polypeptide or its fragments are useful
CC   for treating or preventing neuronal disorder (e.g. akathesia,
CC   Alzheimer's disease, amnesia, dementia or Parkinson's disease),
CC   connective tissue disorder (e.g. cystic fibrosis, myocardial fibrosis,
CC   osteoporosis or Marfan syndrome), or a disorder of cell proliferation
CC   (e.g. cirrhosis, leukaemia, lymphoma, sarcoma or bone marrow cancer).
CC   The polypeptide or its fragments are also useful for screening
CC   libraries of compounds in any of several drug-screening techniques.
XX
SQ   Sequence   314 AA;

  Query Match             38.8%;  Score 537.5;  DB 21;  Length 314;
  Best Local Similarity   45.3%;  Pred. No. 1.7e-45;
  Matches  121;  Conservative   45;  Mismatches   88;  Indels   13;  Gaps    5;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
              |||:|||:    :|     || : |  |:||:| : ||    ||   |   | ||:| ::
Db          3 VGFIGAGQ----LAYRFTAAGILSAHKIIASSP-EMNLPTVSALRKMGVNLTRSNKETVK 57

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELL---PPNTRVLRVL 124
                ::  | |||::| :| |:   |   ||:|| ||||::|::|: |    |  :|:| :
Db         58 HSDVLFLAVKPHIIPFILDEIGADVQARHIVVSCAAGVTISSVEKKLMAFQPAPKVIRCM 117

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
               | | |||||| | | | |    : :||: |: : | | || |  :|  ||||||| |:
Db        118 TNTPVVVQEGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGSGPAY- 176

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
                ||    |:| ||||:|  || :: || ||| |||||   ||| ||: :||:||| ||:
Db        177 -AFMALDADGGVKMGLPRRLAIQLGAQALLGAAKMLLDSEQHPCQLKDNVCSPGGATIHA 235

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
              || || || |:  ::|||:  | :||
Db        236 LHFLESGGFRSLLINAVEASCIRTREL 262


RESULT 15
AAU99322
ID   AAU99322 standard; Protein; 314 AA.
XX
AC   AAU99322;
XX
DT   24-SEP-2002  (first entry)
XX
DE   Human delta 1-pyrroline-5-carboxylate reductase homologue protein.
XX
KW   Human; nootropic; neuroprotective; cytostatic; enzyme; P5CRH;
KW   delta 1-pyrroline-5-carboxylate reductase homologue; proline;
KW   glutamate biosynthesis; collagen; bone; interstitial fluid;
KW   neuronal disorder; Alzheimer's disease; Parkinson's disease;
KW   connective tissue disorder; cystic fibrosis; osteoporosis;
KW   cell proliferation; atherosclerosis; cancer.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 13..15
FT                   /note= "Encoded by TATGCTCTGGCGCGGGGCTTC"
FT   Misc-difference 183..184
FT                   /note= "Encoded by GCATTGGCTGAT"
FT   Region          212..242
FT                   /note= "Delta 1-pyrroline-5-carboxylate reductase
FT                         signature motif"
```

XX
PN US2002081691-A1.
XX
PD 27-JUN-2002.
XX
PF 24-JUL-2001; 2001US-0912717.
XX
PR 18-JUN-1998; 98US-0099676.
PR 05-MAY-2000; 2000US-0565910.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Hillman JL, Corley NC, Baughn MR;
XX
DR WPI; 2002-537133/57.
DR N-PSDB; ABK87030.
XX
PT Novel 1-pyrroline-5-carboxylate reductase homolog polypeptide for
PT treating neuronal disorders, connective tissue disorders, and disorders
PT of cell proliferation -
XX
PS Claim 1; Fig 1; 35pp; English.
XX
CC The invention discloses the purified polypeptide, and polynucleotide
CC encoding it, of human delta 1-pyrroline-5-carboxylate reductase
CC homologue (P5CRH). P5CRH catalyses the NAD(P) dependent oxidation of
CC 1-pyrroline-5-carboxylate to proline, the terminal step in the
CC biosynthesis of proline from glutamate. Proline is a critical substrate
CC for the synthesis of collagen in tissues such as bone and interstitial
CC fluid. The P5CRH and compositions comprising it are useful for the
CC diagnosis and treatment of a condition or disease associated with the
CC expression of P5CRH and prevention of neuronal disorders (e.g.
CC Alzheimer's disease and Parkinson's disease), connective tissue disorders
CC (e.g. cystic fibrosis and osteoporosis), and disorders of cell
CC proliferation (e.g. atherosclerosis and cancer). The polynucleotide may
CC be used as an antisense molecule. Antibodies, which specifically bind
CC P5CRH, may be used for the diagnosis of disorders characterised by
CC expression of P5CRH, or in assays to monitor patients during therapeutic
CC treatment with P5CRH or agonists, antagonists, or inhibitors of P5CRH.
CC Nucleic acid sequences encoding P5CRH may be used to generate
CC hybridisation probes useful in mapping the naturally occurring genomic
CC sequence and may also be used to detect differences in gene sequences
CC among normal, carrier, and affected individuals. P5CRH, its catalytic or
CC immunogenic fragments, or oligopeptides thereof, can be used for
CC screening libraries of compounds in any of a variety of drug screening
CC techniques. The sequence presented is the human delta
CC 1-pyrroline-5-carboxylate reductase homologue protein.
XX
SQ Sequence 314 AA;

Query Match 38.8%; Score 537.5; DB 23; Length 314;
Best Local Similarity 45.3%; Pred. No. 1.7e-45;
Matches 121; Conservative 45; Mismatches 88; Indels 13; Gaps 5;

```
Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
              |||:|||:    :|     || : |  |:||:| : ||    ||   |   | ||:| ::
Db          3 VGFIGAGQ----LAYRFTAAGILSAHKIIASSP-EMNLPTVSALRKMGVNLTRSNKETVK 57

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELL---PPNTRVLRVL 124
                 ::  | |||::| :| |:   |   ||:|| ||||::|::|: |    |  :|:| :
Db         58 HSDVLFLAVKPHIIPFILDEIGADVQARHIVVSCAAGVTISSVEKKLMAFQPAPKVIRCM 117

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
               | | |||||| | | | |    : :||: |: : | | || |  :|  |||||| |:
Db        118 TNTPVVVQEGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGSGPAY- 176

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
                ||    |:| ||||:|  || :: || ||| |||||   ||| ||: :||:||| ||:
Db        177 -AFMALDADGGVKMGLPRRLAIQLGAQALLGAAKMLLDSEQHPCQLKDNVCSPGGATIHA 235

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
               || || || |:  ::|||:  | :||
Db        236 LHFLESGGFRSLLINAVEASCIRTREL 262
```

Search completed: July 30, 2003, 15:27:09
Job time : 87 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 30, 2003, 15:23:05 ; Search time 30 Seconds
(without alignments)
386.439 Million cell updates/sec

Title: US-09-806-536A-14
Perfect score: 1386
Sequence: 1 MAAAEPSPRRVGFVGAGRMA.....AATMSAVEAATCRAKELSRK 274

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	560.5	40.4	315	3	US-09-099-676-3	Sequence 3, Appli
2	560.5	40.4	315	3	US-09-565-910-3	Sequence 3, Appli
3	537.5	38.8	314	3	US-09-099-676-1	Sequence 1, Appli
4	537.5	38.8	314	3	US-09-565-910-1	Sequence 1, Appli
5	345.5	24.9	472	4	US-09-252-991A-17011	Sequence 17011, A
6	327.5	23.6	289	4	US-09-107-532A-7169	Sequence 7169, Ap
7	321	23.2	281	4	US-09-328-352-8170	Sequence 8170, Ap
8	297.5	21.5	304	1	US-08-665-716-2	Sequence 2, Appli
9	269	19.4	282	4	US-09-134-001C-3155	Sequence 3155, Ap
10	179.5	13.0	271	4	US-09-328-352-6802	Sequence 6802, Ap
11	134	9.7	171	4	US-09-252-991A-20711	Sequence 20711, A
12	116	8.4	144	4	US-08-936-165A-445	Sequence 445, App
13	108.5	7.8	409	4	US-09-252-991A-20236	Sequence 20236, A
14	97.5	7.0	1238	4	US-09-252-991A-26363	Sequence 26363, A
15	97	7.0	398	4	US-09-252-991A-20748	Sequence 20748, A
16	96.5	7.0	303	4	US-09-328-352-4879	Sequence 4879, Ap
17	96.5	7.0	516	4	US-09-252-991A-17933	Sequence 17933, A
18	96.5	7.0	811	4	US-09-252-991A-22006	Sequence 22006, A
19	91.5	6.6	436	3	US-08-669-378-4	Sequence 4, Appli
20	91	6.6	280	4	US-09-328-352-4975	Sequence 4975, Ap
21	91	6.6	6396	4	US-09-410-551B-72	Sequence 72, Appl
22	90.5	6.5	851	4	US-09-252-991A-19645	Sequence 19645, A
23	88	6.3	342	4	US-09-364-230-6	Sequence 6, Appli
24	87.5	6.3	221	4	US-09-252-991A-23434	Sequence 23434, A
25	87.5	6.3	537	4	US-09-252-991A-26178	Sequence 26178, A
26	87.5	6.3	565	3	US-09-008-481A-6	Sequence 6, Appli
27	87.5	6.3	565	3	US-09-195-666A-5	Sequence 5, Appli
28	87.5	6.3	565	3	US-09-195-666A-49	Sequence 49, Appl
29	87.5	6.3	565	3	US-09-309-592-6	Sequence 6, Appli
30	87.5	6.3	565	4	US-09-635-705-5	Sequence 5, Appli
31	87.5	6.3	565	4	US-09-635-705-49	Sequence 49, Appl
32	87.5	6.3	565	4	US-09-634-858A-5	Sequence 5, Appli
33	87.5	6.3	565	4	US-09-634-858A-49	Sequence 49, Appl
34	87.5	6.3	565	4	US-08-869-927C-5	Sequence 5, Appli
35	87.5	6.3	565	4	US-08-869-927C-49	Sequence 49, Appl
36	87.5	6.3	762	4	US-09-252-991A-28078	Sequence 28078, A
37	87.5	6.3	1024	4	US-09-562-737-44	Sequence 44, Appl
38	87	6.3	528	3	US-08-928-213B-8	Sequence 8, Appli
39	87	6.3	949	3	US-09-196-387-10	Sequence 10, Appl
40	87	6.3	949	4	US-09-841-835-10	Sequence 10, Appl
41	87	6.3	1327	3	US-09-196-387-2	Sequence 2, Appli
42	87	6.3	1327	4	US-09-841-835-2	Sequence 2, Appli
43	86.5	6.2	418	4	US-09-328-352-5304	Sequence 5304, Ap
44	86.5	6.2	510	4	US-09-252-991A-25949	Sequence 25949, A
45	86.5	6.2	578	4	US-09-252-991A-25113	Sequence 25113, A

ALIGNMENTS

RESULT 1
US-09-099-676-3
; Sequence 3, Application US/09099676
; Patent No. 6100075
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE
; TITLE OF INVENTION: HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/099,676
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0532 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-855-0572
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 315 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 189498
US-09-099-676-3

Query Match 40.4%; Score 560.5; DB 3; Length 315;

```
  Best Local Similarity   46.1%;  Pred. No. 1.6e-53;
  Matches  123;  Conservative   46;  Mismatches   87;  Indels   11;  Gaps    5;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
              |||:|||::  | |:|   || : |  |:||:| | :|    ||   | : |  |:| :|
Db          3 VGFIGAGQL--AFAKGFTAAGVLAAHKIMASSP-DMDLATVSALRKMGVKLTPHNKETVQ 59

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLP---PNTRVLRVL 124
                ::  | |||::| :| |:   :   ||:|| ||||::|::|: |    |  ||:| :
Db         60 HSDVLFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRPAPRVIRCM 119

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
              | | ||:||| | | | |    : :|:: ||   | | || |  :|  |||||||  |:
Db        120 TNTPVVVREGATVYATGTHAQVEDGRLMEQLLSTVGFCTEVEEDLIDAVTGLSGSGPAY- 178

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
               ||:   |:| ||||:|  || |: || ||| ||||||  ||| ||: :| :||| ||:
Db        179 -AFTALDADGGVKMGLPRRLAVRLGAQALLGAAKMLLHSEQHPGQLKDNVSSPGGATIHA 237

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
              || || || |:  ::||||:  | :||
Db        238 LHVLESGGFRSLLINAVEASCIRTREL 264


RESULT 2
US-09-565-910-3
; Sequence 3, Application US/09565910
; Patent No. 6268192
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Baughn, Mariah R.
;    TITLE OF INVENTION:  DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE
;    TITLE OF INVENTION:  HOMOLOG
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/565,910
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/099,676
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cerrone, Michael C
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PF-0532 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-855-0572
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  315 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  189498
US-09-565-910-3


  Query Match           40.4%;  Score 560.5;  DB 3;  Length 315;
  Best Local Similarity   46.1%;  Pred. No. 1.6e-53;
  Matches  123;  Conservative   46;  Mismatches   87;  Indels   11;  Gaps    5;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
              |||:|||::  | |:|   || : |  |:||:| | :|    ||   | : |  |:| :|
Db          3 VGFIGAGQL--AFAKGFTAAGVLAAHKIMASSP-DMDLATVSALRKMGVKLTPHNKETVQ 59

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLP---PNTRVLRVL 124
                ::  | |||::| :| |:   :   ||:|| ||||::|::|: |    |  ||:| :
Db         60 HSDVLFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRPAPRVIRCM 119

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
              | | ||:||| | | | |    : :|:: ||   | | || |  :|  |||||||  |:
Db        120 TNTPVVVREGATVYATGTHAQVEDGRLMEQLLSTVGFCTEVEEDLIDAVTGLSGSGPAY- 178

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
               ||:   |:| ||||:|  || |: || ||| ||||||  ||| ||: :| :||| ||:
Db        179 -AFTALDADGGVKMGLPRRLAVRLGAQALLGAAKMLLHSEQHPGQLKDNVSSPGGATIHA 237

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
              || || || |:  ::||||:  | :||
Db        238 LHVLESGGFRSLLINAVEASCIRTREL 264


RESULT 3
US-09-099-676-1
; Sequence 1, Application US/09099676
; Patent No. 6100075
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Baughn, Mariah R.
;    TITLE OF INVENTION:  DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE
;    TITLE OF INVENTION:  HOMOLOG
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/099,676
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cerrone, Michael C
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PF-0532 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-855-0572
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  314 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  PROSNON01
;      CLONE:  2278458
US-09-099-676-1
```

```
  Query Match             38.8%;  Score 537.5;  DB 3;  Length 314;
  Best Local Similarity   45.3%;  Pred. No. 5.7e-51;
  Matches  121;  Conservative   45;  Mismatches   88;  Indels   13;  Gaps    5;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
              |||:|||:    :|     || : |  |:||:| : ||    ||   |   | ||:| ::
Db          3 VGFIGAGQ----LAYRFTAAGILSAHKIIASSP-EMNLPTVSALRKMGVNLTRSNKETVK 57

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELL---PPNTRVLRVL 124
                ::  | |||::| :| |:   |   ||:|| ||||::|::|: |    |  :|:| :
Db         58 HSDVLFLAVKPHIIPFILDEIGADVQARHIVVSCAAGVTISSVEKKLMAFQPAPKVIRCM 117

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
              | | |||||| | | | |    : :||: |: : | | || |  :|  |||||| |:
Db        118 TNTPVVVQEGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGSGPAY- 176

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
               ||    |:| ||||:|  || :: || ||| |||||   ||| ||: :||:||| ||:
Db        177 -AFMALDADGGVKMGLPRRLAIQLGAQALLGAAKMLLDSEQHPCQLKDNVCSPGGATIHA 235

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
              || || || |:  ::|||:  | :||
Db        236 LHFLESGGFRSLLINAVEASCIRTREL 262


RESULT 4
US-09-565-910-1
; Sequence 1, Application US/09565910
; Patent No. 6268192
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Baughn, Mariah R.
;    TITLE OF INVENTION:  DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE
;    TITLE OF INVENTION:  HOMOLOG
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/565,910
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/099,676
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cerrone, Michael C
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PF-0532 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-855-0572
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  314 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  PROSNON01
;      CLONE:  2278458

US-09-565-910-1

  Query Match             38.8%;  Score 537.5;  DB 3;  Length 314;
  Best Local Similarity   45.3%;  Pred. No. 5.7e-51;
  Matches  121;  Conservative   45;  Mismatches   88;  Indels   13;  Gaps    5;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
              |||:|||:    :|     || : |  |:||:| : ||    ||   |   | ||:| ::
Db          3 VGFIGAGQ----LAYRFTAAGILSAHKIIASSP-EMNLPTVSALRKMGVNLTRSNKETVK 57

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELL---PPNTRVLRVL 124
                ::  | |||::| :| |:   |   ||:|| ||||::|::|: |    |  :|:| :
Db         58 HSDVLFLAVKPHIIPFILDEIGADVQARHIVVSCAAGVTISSVEKKLMAFQPAPKVIRCM 117

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
              | | |||||| | | | |    : :||: |: : | | || |  :|  |||||| |:
Db        118 TNTPVVVQEGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGSGPAY- 176

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
               ||    |:| ||||:|  || :: || ||| |||||   ||| ||: :||:||| ||:
Db        177 -AFMALDADGGVKMGLPRRLAIQLGAQALLGAAKMLLDSEQHPCQLKDNVCSPGGATIHA 235

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
              || || || |:  ::|||:  | :||
Db        236 LHFLESGGFRSLLINAVEASCIRTREL 262


RESULT 5
US-09-252-991A-17011
; Sequence 17011, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 17011
;   LENGTH: 472
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17011

  Query Match             24.9%;  Score 345.5;  DB 4;  Length 472;
  Best Local Similarity   32.2%;  Pred. No. 1.7e-29;
  Matches   87;  Conservative   49;  Mismatches  131;  Indels    3;  Gaps    3;

Qy          7 SPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQA-LGCRTTHSNQEV 65
              |  |: |:||| || ::  || ||  | |  | || |                  || |
Db        201 STPRIAFIGAGNMAASLIGGL-RAQGVPAAQIRASDPGAEQRAKIAGEFAIDVVESNAEA 259

Qy         66 LQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLP 125
              :   :|: : ||  : ||   :|| :  | ::||:|||:  ::||  |     |:| :|
Db        260 VADADVVVLSVKPQAMKAVCQALAPALKPEQLIVSIAAGIPCASLEAWLGQPRPVVRCMP 319

Qy        126 NLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEV-PEAYVDIHTGLSGSGVAFV 184
              | | ::::|| :     | ::: :    || | |    :  || :|  | :|||| |:
Db        320 NTPALLRQGASGLYANAQVSAAQREQAGQLLSAVGIALWLDDEAQIDAVTAVSGSGPAYF 379

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
                  :|: :   |:|:    | |:  || || |:| |     ||:||  | :| |||
Db        380 FLLMQAMTDAGEKLGLSRETASRLTLQTALGAAQMALSSEVEPAELRRRVTSPNGTTEAA 439

Qy        245 LHALEQGGLRAATMSAVEAATCRAKELSRK 274
              : : :  |  |    |: ||: |: ||: :
Db        440 IKSFQANGFEALVEQALNAASQRSAELAEQ 469
```

```
RESULT 6
US-09-107-532A-7169
; Sequence 7169, Application US/09107532A
; Patent No. 6583275
;   GENERAL INFORMATION:
;        APPLICANT: Lynn A Doucette-Stamm  and David Bush
;        TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;                            ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;        NUMBER OF SEQUENCES: 7310
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;             STREET: 100 Beaver Street
;             CITY: Waltham
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02354
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: CD/ROM ISO9660
;             COMPUTER: PC
;             OPERATING SYSTEM: <Unknown>
;             SOFTWARE: ASCII
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/107,532A
;             FILING DATE: 30-Jun-1998
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/085,598
;             FILING DATE: 14 May 1998
;             APPLICATION NUMBER: 60/051571
;             FILING DATE: July 2, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ariniello, Pamela Deneke
;             REGISTRATION NUMBER: 40,489
;             REFERENCE/DOCKET NUMBER: GTC-012
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (781)893-5007
;             TELEFAX: (781)893-8277
;   INFORMATION FOR SEQ ID NO: 7169:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 289 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        HYPOTHETICAL: YES
;        ORIGINAL SOURCE:
;             ORGANISM: Enterococcus faecium
;        FEATURE:
;             NAME/KEY:  misc_feature
;             LOCATION: (B) LOCATION 1...289
;        SEQUENCE DESCRIPTION: SEQ ID NO: 7169:
US-09-107-532A-7169

  Query Match             23.6%;  Score 327.5;  DB 4;  Length 289;
  Best Local Similarity   32.3%;  Pred. No. 7.7e-28;
  Matches   87;  Conservative   51;  Mismatches  124;  Indels    7;  Gaps    5;

Qy         10 RVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQA-LGCRTTHSNQEVLQS 68
              :|||:| | || ||  ||:: | |  : :  |:    |  | |  |     ||||
Db         23 KVGFIGTGNMAQAIIDGLLKKGTVSKEEVFVSSGHFENAQKFAAETGVHACKSNQEVASQ 82

Qy         69 CLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELL-PPNTRVLRVLPNL 127
              :|: | || ::  || |:    : : : :|:|:| ||: ||  |  |   ::||:||:
Db         83 SDIVLLAVKPKIIEKVLEELRE-YSQKVLFISIASGKSLTDLETALEDPQASIVRVMPNV 141

Qy        128 PCVVQEGAIVMARGRHVGSSETKLLQ--HLLEACGRCEEVPEAYVDIHTGLSGSGVAFVC 185
              .  |  || | |  ::  ::| :| |   : || |    : |:      ||:||  |||
Db        142 N--VSVGAGVSAICKNEAATEEQLQQATMIFEAIGSVYPLAESDFSTFIGLAGSSPAFVY 199

Qy        186 AFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGL 245
                :|:    |  |:    |  |||: ::|: :  |   ::| :|   | :|||||: |:
Db        200 MLIDAMGRTGVLNGLSKQTATEIAAKAIMGSCQKFLSSDKNPWELVDQVSSPGGTTVAGV 259

Qy        246 HALEQGGLRAATMSAVEAATCRAKELSRK 274

                ||||: |   | :  :     :  |:  |
Db        260 VALEEAGFIPAIIKGINTTIEKDLEMLNK 288


RESULT 7
US-09-328-352-8170
; Sequence 8170, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTE
;  TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC99-03PA
;  CURRENT APPLICATION NUMBER: US/09/328,352
;  CURRENT FILING DATE: 1999-06-04
;  NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8170
;   LENGTH: 281
;   TYPE: PRT
;   ORGANISM: Acinetobacter baumannii
US-09-328-352-8170

  Query Match             23.2%;  Score 321;  DB 4;  Length 281;
  Best Local Similarity   31.2%;  Pred. No. 3.8e-27;
  Matches   82;  Conservative   55;  Mismatches  122;  Indels    4;  Gaps    4;

Qy         13 FVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTHSNQEVLQSCLLV 72
              |:| | || |:  |||  | :    |  | | ::     |     |  |   :|:  :|
Db         16 FIGGGNMAQALIGGLISRG-LPPTRITVSDPVEQIRQLLQEKEVHVTQDNTAAIQNADVV 74

Qy         73 IFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNLPCVVQ 132
              : | || ||  ||  :  :: :: :::|: ||  : |:  |:  : |::||:|| | :||
Db         75 VLAVKPQVLANVLRPLKGLL-SDKLVISIIAGAEIKTISNLI-DSERIVRVMPNTPALVQ 132

Qy        133 EGAIVMARGRHVGSSETKLLQHLLEACGRCEEV-PEAYVDIHTGLSGSGVAFVCAFSEAL 191
               ||  :     ||:|: :|   :| | |    |  || :|  | :|||| |:     |::
Db        133 TGAHGIYATDVVGASDRELTSQILAATGLTIWVNSEAQIDAVTAVSGSGPAYFFYLMESM 192

Qy        192 AEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHALEQG 251
                    :|:   :|  :  || || |:| :     |::|| :| :| |||   |   ::
Db        193 IRAGKNLGLDEKVATALTLQTALGAAQMAITSSNTPSELRKNVTSPNGTTQAALEVFDRA 252

Qy        252 GLRAATMSAVEAATCRAKELSRK 274
               :     ||: ||  |::||:::
Db        253 QISQNIQSALAAAQKRSQELAQE 275


RESULT 8
US-08-665-716-2
; Sequence 2, Application US/08665716
; Patent No. 5789222
;  GENERAL INFORMATION:
;    APPLICANT:  KELLY, ROSEMARIE
;    APPLICANT:  REGISTER, ELIZABETH A
;    APPLICANT:  MASUREKAR, PRAKASH S
;    TITLE OF INVENTION:  P5C REDUCTASE GENE FROM ZALERION
;    TITLE OF INVENTION:  ARBORICOLA
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MERCK & CO., INC.
;      STREET:  126 E. LINCOLN AVENUE
;      CITY:  RAHWAY
;      STATE:  NEW JERSEY
;      COUNTRY:  US
;      ZIP:  07065-0900
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/665,716
;      FILING DATE:  23-JUN-1995
```

```
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KORSEN, ELLIOTT
;      REGISTRATION NUMBER:  32,705
;      REFERENCE/DOCKET NUMBER:  19453PV
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-594-5493
;      TELEFAX:  908-594-4720
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  304 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-665-716-2

  Query Match             21.5%;  Score 297.5;  DB 1;  Length 304;
  Best Local Similarity   29.3%;  Pred. No. 1.7e-24;
  Matches   88;  Conservative   52;  Mismatches  121;  Indels   39;  Gaps    9;

Qy          3 AAEPSPRRVGFVGAGRMAGAIAQGLIRA-GKVEAQHILASAPTDRNLCHF--QALGCRTT 59
              | | |    :  :| | |  ||  |:: :  :: | :   |: ||        : : |  :
Db          5 AKEGSELTLAVIGCGTMGIAILSGILASLDEIHAPN-SQSSETDETPSKLPTKFIACVRS 63

Qy         60 HSNQEVLQSCL----------------------LVIFATKPHVLPAVLAEVAPVVTTE-H 96
                | ::  |                      :|:   ||::   :| |   |   :
Db         64 PKGAEKIKKALSPYKTPVKIIQSDNVTACREADVVLLGCKPYMAEGILGEEGMVDALKGK 123

Qy         97 ILVSVAAGVSLSTLEELL----PPN------TRVLRVLPNLPCVVQEGAIVMARGR-HVG 145
              :|:|: |||    :   :    | |       :|:| :||    ::|   |:|     :
Db        124 LLISILAGVPAEQIYGYMYGKTPVNPEKEGLCQVVRAMPNTASGIRESMTVIATSSPPLS 183

Qy        146 SSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAFSEALAEGAVKMGMPSSLA 205
              :: : |:  : :  |   ::| | :|  | | |||  ||     ||  :||| ||:| : |
Db        184 ATTSSLITWIFKRIGDVVQLPAATMDASTALCGSGPAFFALILEAAIDGAVAMGLPRAEA 243

Qy        206 HRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHALEQGGLRAATMSAVEAAT 265
              |:||||: | | ::|  |:||| |:  | |||| || ||  ||:||:|     ||  ||
Db        244 QRMAAQTMKGAAGLVL-SGEHPALLKDKVTTPGGCTIGGLMVLEEGGVRGTVARAVREAT 302


RESULT 9
US-09-134-001C-3155
; Sequence 3155, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3155
;   LENGTH: 282
;   TYPE: PRT
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3155

  Query Match             19.4%;  Score 269;  DB 4;  Length 282;
  Best Local Similarity   26.9%;  Pred. No. 2.1e-21;
  Matches   71;  Conservative   54;  Mismatches  137;  Indels    2;  Gaps    2;

Qy         13 FVGAGRMAGAIAQGLIRAGKVEAQHI-LASAPTDRNLCHF-QALGCRTTHSNQEVLQSCL 70
              | ||| || ||  |:| :  : |  | | :   :: |  | : ||   :: :: :|:
Db         16 FYGAGNMAQAIFTGIINSNNLNANDIYLTNKSNEQALKSFAEKLGVNYSYDDEALLKDAD 75

Qy         71 LVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNLPCV 130
                |  ||||    :   :   :| ::  :|: ||:|:  : : |  |  : |::||

Db         76 YVFLGTKPHDFENLANRIREHITNDNRFISIMAGLSIDYIRQQLNTNNPLARIMPNTNAQ 135

Qy        131 VQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAFSEA 190
              |      ::   :         :  |: | |   || | ::   | ::||| ||:    |
Db        136 VGHSVTGISFSNNFDPKSKNEVDELINAFGSVIEVSEEHLHQVTAITGSGSAFLYHVFEQ 195

Qy        191 LAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHALEQ 250
                :   ::|:  :          ::||:||:       :||| :: : ||||  || || |
Db        196 YVKAGTELGLERNQVEESIRNLIIGTSKMIERSDLSMSQLRKNITSKGGTTQAGLDALSQ 255

Qy        251 GGLRAATMSAVEAATCRAKELSRK 274
                : :     : ||  |: ||| |
Db        256 YDIVSMFEDCLGAAVNRSMELSHK 279


RESULT 10
US-09-328-352-6802
; Sequence 6802, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTE
;  TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC99-03PA
;  CURRENT APPLICATION NUMBER: US/09/328,352
;  CURRENT FILING DATE: 1999-06-04
;  NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6802
;   LENGTH: 271
;   TYPE: PRT
;   ORGANISM: Acinetobacter baumannii
US-09-328-352-6802

  Query Match             13.0%;  Score 179.5;  DB 4;  Length 271;
  Best Local Similarity   23.1%;  Pred. No. 1.5e-11;
  Matches   63;  Conservative   51;  Mismatches  124;  Indels   35;  Gaps    5;

Qy         13 FVGAGRMAGAIAQGLIRAG-----------KVEAQHILASAPTDRNLCHFQALGCRTTH 60
              |:|:  :: |:  ||:  |             | | | ||        '
Db         11 FIGSSNLSLALIGGLVLKGFQREKINLIEEVKFENQIIL--------------------K 50

Qy         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120
                |  ::   :|:    |  | |:|| :   :  :  :||: |||::  |   :  : ::
Db         51 QKQHEVKKADIVVLLLDPKDLKAILAPLKKWL-ADKTIVSMMAGVNIQQLMS-ITGSKKI 108

Qy        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEV-PEAYVDIHTGLSGS 179
              :||: | | :   |  |:    ::   : :::: :  | ||      |:  |    ||||
Db        109 IRVISNPPVLTYTGTHVLIGSDYLEPLDKEVIETIYSATGRTYWANSESQSDAIIALSGS 168

Qy        180 GVAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGG 239
              | |:     ::: :  | ||:    |  :  |   |  :|:      |::|   |    |
Db        169 GPAYFFYILDSMVKTGVSMGLDKQFALDLILQAASGAVEMVRKSNVQPSELCGKVTLANG 228

Qy        240 TTIYGLHALEQGGLRAATMSAVEAATCRAKELS 272
                |   |   | | |      |::||  |:||::
Db        229 ITESALRMFELGNLSDDIRLALKAAYHRSKEIN 261


RESULT 11
US-09-252-991A-20711
; Sequence 20711, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
```

```
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 20711
;   LENGTH: 171
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20711

  Query Match             9.7%;  Score 134;  DB 4;  Length 171;
  Best Local Similarity   33.7%;  Pred. No. 8e-07;
  Matches   29;  Conservative   17;  Mismatches   40;  Indels    0;  Gaps    0;

Qy        189 EALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHAL 248
              :|: :   |:|:    | |:  || || |:| |      ||:||  | :| |||   : :
Db         83 QAMTDAGEKLGLSRETASRLTLQTALGAAQMALSSEVEPAELRRRVTSPNGTTEAAIKSF 142

Qy        249 EQGGLRAATMSAVEAATCRAKELSRK 274
              :  |  |    |: ||: |: ||: :
Db        143 QANGFEALVEQALNAASQRSAELAEQ 168


RESULT 12
US-08-936-165A-445
; Sequence 445, Application US/08936165A
; Patent No. 6348582
;  GENERAL INFORMATION:
;    APPLICANT:  Black, Michael
;    APPLICANT:  Burnham, Martin
;    APPLICANT:  Hodgson, John
;    APPLICANT:  Knowles, David
;    APPLICANT:  Lonetto, Michael
;    APPLICANT:  Nicholas, Richard
;    APPLICANT:  Pratt, Julie
;    APPLICANT:  Reichard, Richard
;    APPLICANT:  Rosenberg, Martin
;    APPLICANT:  Ward, Judith
;    TITLE OF INVENTION:  No. 6348582el Prokaryotic Polynucleotides,
;    TITLE OF INVENTION:  Polypeptides and Their Uses
;    NUMBER OF SEQUENCES:  534
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19406-0939
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/936,165A
;      FILING DATE:  24-SEP-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/027,032
;      FILING DATE:  24-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gimmi, Edward R
;      REGISTRATION NUMBER:  38,891
;      REFERENCE/DOCKET NUMBER:  P50549
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-4478
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  445:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  144 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Protein
US-08-936-165A-445

  Query Match             8.4%;  Score 116;  DB 4;  Length 144;
  Best Local Similarity   26.8%;  Pred. No. 6e-05;
  Matches   30;  Conservative   24;  Mismatches   56;  Indels    2;  Gaps    2;

Qy         16 AGRMAGAIAQGLIRAGKVEAQHI-LASAPTDRNLCHF-QALGCRTTHSNQEVLQSCLLVI 73
              || || ||  |:| :  ::|  | | :   :: |  | : ||   :: :  :|:    |
Db         12 AGNMAQAIFTGIINSSNLDANDIYLTNKSNEQALKAFAEKLGVNYSYDDATLLKDADYVF 71

Qy         74 FATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLP 125
                ||||   |:   : | :| :   :|: ||:    : : |       |::|
Db         72 LGTKPHDFDALATRIKPHITKDXCFMSIMAGIPTDYIXQQLECQNPXARIMP 123


RESULT 13
US-09-252-991A-20236
; Sequence 20236, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 20236
;   LENGTH: 409
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20236

  Query Match             7.8%;  Score 108.5;  DB 4;  Length 409;
  Best Local Similarity   26.3%;  Pred. No. 0.002;
  Matches   74;  Conservative   30;  Mismatches  102;  Indels   75;  Gaps   17;

Qy          4 AEPSPRRV--GFVGAGRMA-GAIAQGLIRAGKVEAQHILASAPTDRNLCHFQA-LGCRTT 59
              | | ||    |  | | :  |  | |:||      ::  |:|   : |   || : ||
Db         90 ALPGPRATLEGHPGGGSVPWGEYATGVIR------RYPAATADNRKRLFALQAHVPCRKA 143

Qy         60 HSNQEVLQSCLLVIFATKPHVLPAVLAEVAPV-VTTEHILVSVAAGVSLSTLEELLPPNT 118
                        |:          |||||    |  | |:|  | |   | |  :: :| |
Db        144 TRGWPF---CMATNATGAIRCLPAVLLHRHPAQVDTDHHRVVVGGPVRLGVVQAVLLPRG 200

Qy        119 RVLRVL---PNLPCV---------------VQEGAIVMA--RGRHVG--SSETKLLQHLL 156
              | |:   | |  |                | || : :|   | |||   |  :| :| :
Db        201 LELDVVVQHPGLVGVDVPLHRAAYLARQAGVGEGVVKVAGLPGEHVGIALSCARLTRHQV 260

Qy        157 EACGRCEEVPEAYVDIHTGLSGSGVAFVCAFSEALAEGAVK-MGMPSSLAHRIAAQTLLG 215
              |  :       |:|    | |:       |   : |:: :|     | |||  |
Db        261 GVAGAAD-------DVHRVFLGIGI-------EVADDQAIRLLG-----ARRIARQP--- 298

Qy        216 TAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHALEQGGLRAA 256
                 :|:|   |:       ||   :    ||   |:| |
Db        299 -----VHQGLGGAR-------PGQVAV----ALAVAGVRVA 323


RESULT 14
US-09-252-991A-26363
; Sequence 26363, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
```

```
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 26363
;   LENGTH: 1238
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26363

  Query Match             7.0%;  Score 97.5;  DB 4;  Length 1238;
  Best Local Similarity   25.0%;  Pred. No. 0.17;
  Matches   58;  Conservative   24;  Mismatches   73;  Indels   77;  Gaps   12;

Qy         23 IAQGLIRAGKVE-----AQHILASAPTDRNLCHFQALGCRTTHSNQEVLQSCLLVIFATK 77
              :||||  ||:        ||:|  ||  | : | |
Db        981 LAQGLGHAGRQVRRRGFLQHLLV-APLHRAVAHAQG----------------------- 1015

Qy         78 PHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNLPCVVQEGAIV 137
                  :||  |  ||: : ||      ||: ||  : |:  |:  ||
Db        1016 --------DDVAGAV-AEHLYLQVA-----GTLDVLLDEHARIAEVV--LP--------- 1050

Qy        138 MARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHT---GLSGSGVAFVCAFSEALAEG 194
               |  |  | || :                |: |  |          || :   :| : |
Db        1051 QALDRVEGISELR------------RRAAHAHADTTTTGGAFQHHRVADLLGCAEGIGEA 1098

Qy        195 AVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPA--QLRSDVCTPGGTTIYG 244
              ::|   :| || ||  | |     : | :||   : |:|   |||   :|
Db        1099 VEQLG---ALQHRYAA--LFGQGAGGVLEAKHPQLFRRRADEGDPGGLAGFG 1145


RESULT 15
US-09-252-991A-20748
; Sequence 20748, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 20748
;   LENGTH: 398
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20748

  Query Match             7.0%;  Score 97;  DB 4;  Length 398;
  Best Local Similarity   23.7%;  Pred. No. 0.035;
  Matches   81;  Conservative   34;  Mismatches  121;  Indels  106;  Gaps   18;

Qy          2 AAAEPSPRRVGFVGAGRMAGAIAQGLI-------------------------RAGKVEAQ 36
              | : ||| |  |       |   |:||                          |    ||
Db         40 AMSLPSPWRADFPAFSAFA---AEGLTYLDSAATAQKPQAVLDALNGYYLGGAANVHRAQ 96

Qy         37 HILASAPTDRNLCHFQALGCRTTH-----SNQEVL------QSCLLVIFATKPHVLPAVL 85
              |:    | :|    |:|   |  |     :  |||      ::  || :  : |  |
Db         97 HV----PGERATRAFEAARSRVAHWLHAGNPAEVLFTRGTTEAINLVAYGLERHFRPG-- 150

Qy         86 AEVAPVVTTEHILVS-VAAGVSLSTLEELLPPNTRVLRVLPNLPCVVQEGAIVMARGRHV 144
                    : :||| :    :|   ::|      ||||||    : : | | : : ||:
Db        151 ---------DELLVSALEHHANLLPWQQLALRRGLVLRVLP----LDERGVIDLEQARHI 197

Qy        145 GSSETKLL--QHLLEACGRCEEVPE---------------AYVDIH--TGLSGSGVAFV 184
                |:||    |    |  : | |                |   :|    | | |  |
Db        198 IGERTRLLAISQLSNVLGTWQPVVELIQLARERGAWTLVDGAQGSVHGRHDLPGLGCDFY 257

Qy        185 CAFSEALAEG----AVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRS-----DVC 235
              |||      |      |  | | :|      |     |  :|: | | | |    :      :
Db        258 -AFSGHKLYGPDGIGVLWGRPQALEQ--LAHWQFG-GEMVRHTGFHEASFHAAPLGFEAG 313

Qy        236 TPG-------GTTIYGLHALE-------QGGLRAATMSAVEA 263
              ||        | ||  |  |:       :| | |  :: : |
Db        314 TPAVSAAIGLGATIDWLATLDEAEVAAHEGALHARLLAGLLA 355


Search completed: July 30, 2003, 15:28:42
Job time : 31 secs
```

OM protein - protein search, using sw model

Run on: July 30, 2003, 15:23:05 ; Search time 98 Seconds
(without alignments)
721.494 Million cell updates/sec

Title: US-09-806-536A-14
Perfect score: 1386
Sequence: 1 MAAAEPSPRRVGFVGAGRMA.....AATMSAVEAATCRAKELSRK 274

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1386	100.0	274	4	Q9H896	Q9h896 homo sapien
2	1381	99.6	274	4	Q96HX4	Q96hx4 homo sapien
3	1235	89.1	244	4	Q8N3N9	Q8n3n9 homo sapien
4	1163	83.9	274	11	Q9DCC4	Q9dcc4 mus musculu
5	1160	83.7	274	11	Q8R0P9	Q8r0p9 mus musculu
6	921	66.5	248	11	Q9D0X2	Q9d0x2 mus musculu
7	613	44.2	279	5	Q21544	Q21544 caenorhabdi
8	585	42.2	270	5	Q8MUT0	Q8mut0 tigriopus c
9	577.5	41.7	320	4	Q96C36	Q96c36 homo sapien
10	571.5	41.2	320	11	Q922Q4	Q922q4 mus musculu
11	560.5	40.4	273	5	Q9VEJ3	Q9vej3 drosophila
12	560.5	40.4	309	11	Q922W5	Q922w5 mus musculu
13	555.5	40.1	274	2	Q8GE04	Q8ge04 heliobacill
14	524.5	37.8	270	17	Q8PYP4	Q8pyp4 methanosarc
15	505.5	36.5	280	5	Q9V3F8	Q9v3f8 drosophila
16	503	36.3	274	10	Q9AYM3	Q9aym3 vigna ungui
17	493.5	35.6	280	5	O96643	O96643 drosophila
18	480	34.6	270	16	Q97E64	Q97e64 clostridium
19	477	34.4	319	4	Q9Y5J4	Q9y5j4 homo sapien
20	469	33.8	266	16	Q8Y9X2	Q8y9x2 listeria mo
21	458	33.0	266	16	Q92EP4	Q92ep4 listeria in
22	457	33.0	284	10	Q8GT01	Q8gt01 oryza sativ
23	455.5	32.9	270	16	Q8YZH2	Q8yzh2 anabaena sp
24	438.5	31.6	266	2	O87725	O87725 clostridium
25	434.5	31.3	266	16	Q8XH52	Q8xh52 clostridium
26	430	31.0	266	16	Q8KG91	Q8kg91 chlorobium
27	429	31.0	264	16	Q8RA68	Q8ra68 thermoanaer
28	426.5	30.8	272	4	Q9HBQ4	Q9hbq4 homo sapien
29	425.5	30.7	305	16	Q8FKE0	Q8fke0 escherichia
30	423.5	30.6	269	16	Q8XEB1	Q8xeb1 escherichia
31	421.5	30.4	269	16	Q8ZRF0	Q8zrf0 salmonella
32	420.5	30.3	284	16	Q9X8G1	Q9x8g1 streptomyce
33	419.5	30.3	269	16	Q8Z8Z2	Q8z8z2 salmonella
34	418.5	30.2	262	5	Q8IDC6	Q8idc6 plasmodium
35	416.5	30.1	272	2	Q9K308	Q9k308 bacillus ce
36	405	29.2	261	17	Q8U084	Q8u084 pyrococcus
37	404	29.1	174	5	Q8IN96	Q8in96 drosophila
38	395	28.5	265	16	Q97R93	Q97r93 streptococc
39	394	28.4	283	3	Q96WX7	Q96wx7 emericella
40	385	27.8	265	16	Q8DQ59	Q8dq59 streptococc
41	371.5	26.8	256	16	Q8DS61	Q8ds61 streptococc
42	369	26.6	158	5	Q8MUT1	Q8mut1 tigriopus c
43	367	26.5	254	5	Q9NGS4	Q9ngs4 leishmania
44	364.5	26.3	314	10	Q8S488	Q8s488 phytophthor
45	363.5	26.2	273	16	Q8DJK6	Q8djk6 synechococc

ALIGNMENTS

RESULT 1
Q9H896
ID Q9H896 PRELIMINARY; PRT; 274 AA.
AC Q9H896;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ13852.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thyroid;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AK023914; BAB14721.1; -.
DR InterPro; IPR000304; P5CR.
DR Pfam; PF01089; P5CR; 1.
DR TIGRFAMs; TIGR00112; proC; 1.
DR PROSITE; PS00521; P5CR; 1.
KW Hypothetical protein.
SQ SEQUENCE 274 AA; 28663 MW; 846FDEC603F3B548 CRC64;

Query Match 100.0%; Score 1386; DB 4; Length 274;
Best Local Similarity 100.0%; Pred. No. 3.9e-103;
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60

Qy 61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120

Qy        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180

Qy        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240

Qy        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274
                 ||||||||||||||||||||||||||||||||||
Db        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274


RESULT 2
Q96HX4
ID   Q96HX4      PRELIMINARY;      PRT;   274 AA.
AC   Q96HX4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RA   Strausberg R.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC007993; AAH07993.1; -.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   274 AA;  28649 MW;  846FD9B60183B048 CRC64;

  Query Match             99.6%;  Score 1381;  DB 4;  Length 274;
  Best Local Similarity   99.6%;  Pred. No. 9.7e-103;
  Matches  273;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60

Qy         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120

Qy        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180
                 |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETNLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180

Qy        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240

Qy        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274
                 ||||||||||||||||||||||||||||||||||
Db        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274


RESULT 3
Q8N3N9
ID   Q8N3N9      PRELIMINARY;      PRT;   244 AA.
AC   Q8N3N9;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   DKFZP761H0716.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Amygdala;
RA   Ansorge W., Wirkner U., Mewes H.W., Weil B., Wiemann S.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL833857; CAD38716.1; -.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Hypothetical protein.
FT   NON_TER       1       1
SQ   SEQUENCE   244 AA;  25684 MW;  421656FB0FBA0741 CRC64;

  Query Match             89.1%;  Score 1235;  DB 4;  Length 244;
  Best Local Similarity   99.6%;  Pred. No. 4.1e-91;
  Matches  243;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         31 GKVEAQHILASAPTDRNLCHFQALGCRTTHSNQEVLQSCLLVIFATKPHVLPAVLAEVAP 90
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GKVEAQHILASAPTDRNLCHFQALGCRTTHSNQEVLQSCLLVIFATKPHVLPAVLAEVAP 60

Qy         91 VVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNLPCVVQEGAIVMARGRHVGSSETK 150
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNLPCVVQEGAIVMARGRHVGSSETN 120

Qy        151 LLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAFSEALAEGAVKMGMPSSLAHRIAA 210
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAFSEALAEGAVKMGMPSSLAHRIAA 180

Qy        211 QTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHALEQGGLRAATMSAVEAATCRAKE 270
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 QTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHALEQGGLRAATMSAVEAATCRAKE 240

Qy        271 LSRK 274
                 ||||
Db        241 LSRK 244


RESULT 4
Q9DCC4
ID   Q9DCC4      PRELIMINARY;      PRT;   274 AA.
AC   Q9DCC4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   1110058B13Rik protein.
GN   1110058B13RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Kidney;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
```

```
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK002912; BAB22451.1; -.
DR   MGD; MGI:1913444; 1110058B13Rik.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
SQ   SEQUENCE   274 AA;  28694 MW;  FE3892C01C6068A5 CRC64;

  Query Match             83.9%;  Score 1163;  DB 11;  Length 274;
  Best Local Similarity   82.5%;  Pred. No. 2.7e-85;
  Matches  226;  Conservative   23;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60
              |||    |||||||||||||| |||:|||:||||||: :||||||| |||||:||||:|||
Db          1 MAATMSEPRRVGFVGAGRMAEAIARGLIQAGKVEAKQVLASAPTDNNLCHFRALGCQTTH 60

Qy         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120
              || |||||| |||||||| ||| |||||||:||||||:||||||:||||:| ||||||||
Db         61 SNHEVLQSCPLVIFATKPQVLPTVLAEVAPIVTTEHIIVSVAAGISLSTMEGLLPPNTRV 120

Qy        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180
              ||| |||||||||||:||||| | |: : :|||:||||||:| ||||:||||||||||||
Db        121 LRVSPNLPCVVQEGAMVMARGHHAGNDDAELLQNLLEACGQCIEVPESYVDIHTGLSGSG 180

Qy        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240
              ||||| |||||||||:|||||| |||||||||||||||||  ||:||||||:|| || ||
Db        181 VAFVCTFSEALAEGAIKMGMPSGLAHRIAAQTLLGTAKMLQQEGKHPAQLRTDVLTPAGT 240

Qy        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274
              ||:||||||:|| |||||||||||||||||:|
Db        241 TIHGLHALERGGFRAATMSAVEAATCRAKELSKK 274


RESULT 5
Q8R0P9
ID   Q8R0P9      PRELIMINARY;      PRT;   274 AA.
AC   Q8R0P9;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   RIKEN cDNA 1110058B13 gene.
GN   1110058B13RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Colon;
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC026536; AAH26536.1; -.
DR   MGD; MGI:1913444; 1110058B13Rik.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
SQ   SEQUENCE   274 AA;  28721 MW;  E01FB7133B45BD7C CRC64;

  Query Match             83.7%;  Score 1160;  DB 11;  Length 274;
  Best Local Similarity   82.1%;  Pred. No. 4.8e-85;
  Matches  225;  Conservative   24;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60
              |||    |||||||||||||| |||:|||:||||||: :||||||| |||||:||||:|||
Db          1 MAATMSEPRRVGFVGAGRMAEAIARGLIQAGKVEAKQVLASAPTDNNLCHFRALGCQTTH 60

Qy         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120
              || |||||| |||||||| ||| |||||||:||||||:||||||:||||:| ||||||||
Db         61 SNHEVLQNCPLVIFATKPQVLPTVLAEVAPIVTTEHIIVSVAAGISLSTMEGLLPPNTRV 120

Qy        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180
              ||| |||||||||||:||||| | |: : :|||:||||||:| ||||:||||||||||||
Db        121 LRVSPNLPCVVQEGAMVMARGHHAGNDDAELLQNLLEACGQCIEVPESYVDIHTGLSGSG 180

Qy        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240
              ||||| |||||||||:|||||| |||||||||||||||||  ||:||||||:|| || ||
Db        181 VAFVCTFSEALAEGAIKMGMPSGLAHRIAAQTLLGTAKMLQQEGKHPAQLRTDVLTPAGT 240

Qy        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274
              ||:||||||:|| |||||||||||||||||:|
Db        241 TIHGLHALERGGFRAATMSAVEAATCRAKELSKK 274


RESULT 6
Q9D0X2
ID   Q9D0X2      PRELIMINARY;      PRT;   248 AA.
AC   Q9D0X2;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   1110058B13Rik protein.
GN   1110058B13RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK004291; BAB23252.1; -.
DR   MGD; MGI:1913444; 1110058B13Rik.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
SQ   SEQUENCE   248 AA;  26170 MW;  4AFA896396305A34 CRC64;

  Query Match             66.5%;  Score 921;  DB 11;  Length 248;
  Best Local Similarity   73.0%;  Pred. No. 5.7e-66;
  Matches  181;  Conservative   27;  Mismatches   36;  Indels    4;  Gaps    2;

Qy          1 MAAAEPSPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTH 60
              |||    |||||||||||||| |||:|||:||||||: :||||||| |||||:||||:| |
Db          1 MAATMSEPRRVGFVGAGRMAEAIARGLIQAGKVEAKQVLASAPTDNNLCHFRALGCQTNH 60

Qy         61 SNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRV 120
              || |||||| |||||||| ||| |||||||:: ||||:||||||:|  | | |||||||
Db         61 SNHEVLQNCPLVIFATKPQVLPTVLAEVAPIINTEHIIVSVAAGISHGTKEGLLPPNTRA 120
```

```
Qy        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180
              ||| ||| |||||||:||||| | |: : :|:|:||||||:| :|||:|||||||||||
Db        121 LRVSPNLTCVVQEGAMVMARGHHAGNDDAELIQNLLEACGQCIKVPESYVDIHTGLSGSG 180

Qy        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240
              ||||| ||| |||||:|:|||| |||||||||||||||  ||:|||||  | |:
Db        181 VAFVCTFSEGLAEGAIKIGMPSGLAHRIAAQTLLGTAKMLQQEGKHPAQL-PDRCS---H 236

Qy        241 TIYGLHAL 248
              | :  |::
Db        237 TCWNHHSM 244


RESULT 7
Q21544
ID   Q21544      PRELIMINARY;      PRT;   279 AA.
AC   Q21544;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   M153.1 protein.
GN   M153.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Matthews P.;
RL   Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998)..
DR   EMBL; Z67995; CAA91943.1; -.
DR   WormPep; M153.1; CE03511.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
SQ   SEQUENCE   279 AA;  29165 MW;  CE8CAD4BC4EB20E6 CRC64;

  Query Match             44.2%;  Score 613;  DB 5;  Length 279;
  Best Local Similarity   46.1%;  Pred. No. 3e-41;
  Matches  123;  Conservative   56;  Mismatches   86;  Indels    2;  Gaps    1;

Qy         10 RVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAP--TDRNLCHFQALGCRTTHSNQEVLQ 67
              ::||:|||:|| |:|:||| :|:: | :|:||:|   :  |   :|||  ||| | ||:|
Db          2 KIGFIGAGKMAQALARGLINSGRITADNIIASSPKRDEVFLDQCKALGLNTTHDNAEVVQ 61

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNL 127
                 :|  | ||  :  | :|:|| :: ||::||:| |:::  :| |||  :||:||:||
Db         62 KSDVVFLAVKPVHVSKVASEIAPALSKEHLVVSIALGITIRNIESLLPTKSRVVRVMPNT 121

Qy        128 PCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAF 187
              | ||: ||   | |      : : :: ||   |   |||| ::|  |||||| ::: |
Db        122 PSVVRAGASAFAMGSACRDGDAETVEKLLSTVGFAVEVPEIHIDPVTGLSGSGPSYMFAV 181

Qy        188 SEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHA 247
               | ||:| ||:|:|  || ::|| |||| |||:|  | |||||: || :| |:::||:|
Db        182 IEGLADGGVKVGLPRDLALKLAAYTLLGAAKMVLETGIHPAQLKDDVQSPAGSSVYGMHK 241

Qy        248 LEQGGLRAATMSAVEAATCRAKELSRK 274
              || |||:   | |||||| |::    |
Db        242 LESGGLKGVLMDAVEAATNRSRATGDK 268


RESULT 8
Q8MUT0
ID   Q8MUT0      PRELIMINARY;      PRT;   270 AA.
AC   Q8MUT0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyrroline-5-carboxylase reductase.
GN   P5CR.
OS   Tigriopus californicus.
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Maxillopoda; Copepoda;
OC   Harpacticoida; Harpactidae; Tigriopus.
OX   NCBI_TaxID=6832;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SD;
RA   Willett C.S., Burton R.S.;
RT   "Proline biosynthesis genes and their regulation under salinity stress
RT   in the euryhaline copepod Tigriopus californicus.";
RL   Comp. Biochem. Physiol. B, Comp. Biochem. 132:739-750(2002).
DR   EMBL; AF512515; AAM48241.1; -.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
SQ   SEQUENCE   270 AA;  28482 MW;  4CD88DE990EFBB82 CRC64;

  Query Match             42.2%;  Score 585;  DB 5;  Length 270;
  Best Local Similarity   43.6%;  Pred. No. 5e-39;
  Matches  116;  Conservative   63;  Mismatches   83;  Indels    4;  Gaps    4;

Qy         10 RVGFVGAGRMAGAIAQGLIRAGKVE-AQHILASAP-TDRNLC-HFQALGCRTTHSNQEVL 66
              ||||:|||:|| |:||| |:   |: | |: || |  |::|  | : :||:| | |:|::
Db          6 RVGFIGAGKMASALAQGFIKGQAVKCASHVSASCPIADKHLLEHMKQMGCQTFHCNKELM 65

Qy         67 QSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPN 126
                 :|| | || ::| || ::  :| :: ::||:|||| |: |:  | |:::::|::||
Db         66 DCSDVVILAVKPQIMPLVLPDIKSLVASK-LMVSIAAGVQLAQLQADLNPDSKIVRLMPN 124

Qy        127 LPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCA 186
              ||:|:||  |   | |  :   :: |    |: :|| |: :|  ||:|||| |::
Db        125 TPCIVREGVSVFCSGPQVTPQDNAIVKQLFGTVGQIDEVKESMIDAVTGVSGSGPAYMYL 184

Qy        187 FSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLH 246
               ||:|:| || |:|  ||:::||||::|  :|:|    ||| |: :|||:||||||  |
Db        185 IIEAMADGGVKQGLPRDLAYKLAAQTMVGAGRMVLDTHTHPAVLKDEVCSPGGTTITALD 244

Qy        247 ALEQGGLRAATMSAVEAATCRAKELS 272
              ||: ||| : | |||||| |  ::|
Db        245 MLEKSGLRTSIMQAVEAATKRCTQIS 270


RESULT 9
Q96C36
ID   Q96C36      PRELIMINARY;      PRT;   320 AA.
AC   Q96C36;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to pyrroline 5-carboxylate reductase isoform (Hypothetical
DE   protein).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Cervix, and Skin;
RA   Strausberg R.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC014868; AAH14868.1; -.
DR   EMBL; BC020553; AAH20553.1; -.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Hypothetical protein.
```

```
SQ   SEQUENCE   320 AA;  33637 MW;  9348455A64CCE722 CRC64;

  Query Match             41.7%;  Score 577.5;  DB 4;  Length 320;
  Best Local Similarity   46.4%;  Pred. No. 2.4e-38;
  Matches  124;  Conservative   48;  Mismatches   88;  Indels    7;  Gaps    3;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
              |||:|||::| |:|:|   || : |  |:||:| : ||    ||   |   | ||:| ::
Db          3 VGFIGAGQLAYALARGFTAAGILSAHKIIASSP-EMNLPTVSALRKMGVNLTRSNKETVK 61

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELL---PPNTRVLRVL 124
                ::  | |||::| :| |:   |   ||:|| ||||::|::|: |    |  :|:| :
Db         62 HSDVLFLAVKPHIIPFILDEIGADVQARHIVVSCAAGVTISSVEKKLMAFQPAPKVIRCM 121

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
               | | ||||||  | | | |    : :||: |: : | | || |  :|  |||||| |:
Db        122 TNTPVVVQEGATVYATGTHALVEDGQLLEQLMSSVGFCTEVEEDLIDAVTGLSGSGPAYA 181

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
                :|||:| ||||:|  || :: || ||| |||||   ||| ||: :||:||| ||:
Db        182 FMALDALADGGVKMGLPRRLAIQLGAQALLGAAKMLLDSEQHPCQLKDNVCSPGGATIHA 241

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
              || || || |:  ::|||||:  | :||
Db        242 LHFLESGGFRSLLINAVEASCIRTREL 268


RESULT 10
Q922Q4
ID   Q922Q4      PRELIMINARY;      PRT;   320 AA.
AC   Q922Q4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Similar to pyrroline 5-carboxylate reductase isoform.
GN   1810018M05RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC006882; AAH06882.1; -.
DR   MGD; MGI:1916301; 1810018M05Rik.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
SQ   SEQUENCE   320 AA;  33659 MW;  64DD7F96FB8C4BC5 CRC64;

  Query Match             41.2%;  Score 571.5;  DB 11;  Length 320;
  Best Local Similarity   46.1%;  Pred. No. 7.3e-38;
  Matches  123;  Conservative   48;  Mismatches   89;  Indels    7;  Gaps    3;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
              |||:|||::| |:|:|   || : |  |:||:| | :|    ||   |   | ||:: ::
Db          3 VGFIGAGQLACALARGFTAAGVLSAHKIIASSP-DMDLPTVSALRRMGVNLTRSNKDTVR 61

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELL---PPNTRVLRVL 124
                ::  | |||::| :| |:   |   ||:|| ||||::|::|: |    |  :|:| :
Db         62 HSDVLFLAVKPHIIPFILDEIGADVQERHIVVSCAAGVTISSVEKKLMAFQPAPKVIRCM 121

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
               | | ||:||| | | | |    : |||: |: : | | || |  :|  |||||| |:
Db        122 TNTPVVVREGATVYATGTHALVEDGKLLEQLMSSVGFCTEVEEDLIDAITGLSGSGPAYA 181

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
                :|||:| ||||:|  || |: || ||| |||||    || ||: :||:||| ||:
Db        182 FMALDALADGGVKMGVPRRLAVRLGAQALLGAAKMLLDSEDHPGQLKDNVCSPGGATIHA 241

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
              || || || |:  ::|||||:  | :||
Db        242 LHFLESGGFRSLLINAVEASCIRTREL 268


RESULT 11
Q9VEJ3
ID   Q9VEJ3      PRELIMINARY;      PRT;   273 AA.
AC   Q9VEJ3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   CG5840 protein (RE62767p).
GN   CG5840.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003718; AAF55428.1; -.
DR   EMBL; AY071558; AAL49180.1; -.
DR   FlyBase; FBgn0038516; CG5840.
```

```
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
SQ   SEQUENCE   273 AA;  28099 MW;  563754C62ACE77AB CRC64;

  Query Match             40.4%;  Score 560.5;  DB 5;  Length 273;
  Best Local Similarity   44.2%;  Pred. No. 4.6e-37;
  Matches  118;  Conservative   50;  Mismatches   96;  Indels    3;  Gaps    3;

Qy         10 RVGFVGAGRMAGAIAQGLIRAGKVEAQHILASA-PTDR-NLCHFQALGCRTTHSNQEVLQ 67
              ::||:| | || |:|:| : ||  :   ::||  | |: :|  ||:||  |   |  |:|
Db          6 KIGFLGGGNMAKALAKGFLAAGLAKPNTLIASVHPADKLSLQSFQSLGVETVIKNAPVVQ 65

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNL 127
                :|  : || |:|:||:|: | :::  : :||| |::|||:|  | |  ||:||:|||
Db         66 QSDVVFVSVKPQVVPSVLSEIQP-LSSGKLFLSVAMGITLSTIESSLSPQARVIRVMPNL 124

Qy        128 PCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAF 187
              | ||  |  |  ||      ::  : | ||:: | || | |: :|: | |||||  |:|
Db        125 PAVVCSGCSVFVRGSKATDADADITQKLLQSVGTCEPVDESQLDVVTALSGSGPAYVFVM 184

Qy        188 SEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHA 247
               ||||:||| |||| ||:|:|:||:||   |:   | || ||:  | :| |:|   |
Db        185 IEALADGAVHMGMPRDLAYRLASQTVLGAGHMVRDSGMHPGQLKDGVTSPAGSTAAALRQ 244

Qy        248 LEQGGLRAATMSAVEAATCRAKELSRK 274
              ||  | |||   ||| || | :::| |
Db        245 LELSGFRAAVSGAVEQATLRCRQISGK 271


RESULT 12
Q922W5
ID   Q922W5      PRELIMINARY;      PRT;   309 AA.
AC   Q922W5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Similar to pyrroline-5-carboxylate reductase 1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC006727; AAH06727.1; -.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
SQ   SEQUENCE   309 AA;  32373 MW;  A3CD24AACDD53DEF CRC64;

  Query Match             40.4%;  Score 560.5;  DB 11;  Length 309;
  Best Local Similarity   43.2%;  Pred. No. 5.3e-37;
  Matches  115;  Conservative   54;  Mismatches   92;  Indels    5;  Gaps    2;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAP--TDRNLCHFQALGCRTTHSNQEVLQS 68
              |||:|||::| |:|:|   || : |  |:|||:|      :   : :|   |  |:| ::
Db          3 VGFIGAGQLAFALAKGFTAAGVLAAHKIMASSPDMDQATVSALRKIGVNLTPHNKETVRH 62

Qy         69 CLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELL---PPNTRVLRVLP 125
              ::  | |||::| :| |:   :   ||:|| ||||:::::|: |    |  :|:| :
Db         63 SDVLFLAVKPHIIPFILDEIGANIEDRHIVVSCAAGVTINSIEKKLTAFQPAPKVIRCMT 122

Qy        126 NLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVC 185
              | | ||:||  | | | |    : :|:: |: : | | || |  :|  |||||||  |:
Db        123 NTPVVVREGVTVYATGTHAQVEDGRLVEQLMGSVGFCTEVEEDLIDAVTGLSGSGPAYAF 182

Qy        186 AFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGL 245
               :|||:| ||||:|  || |: || ||| |||||   |||:||: :||:||| ||: |
Db        183 TALDALADGGVKMGLPRRLAVRLGAQALLGAAKMLLDSEQHPSQLKDNVCSPGGATIHAL 242
```

```
Qy        246 HALEQGGLRAATMSAVEAATCRAKEL 271
              | || || |:  ::||||:  | :||
Db        243 HVLESGGFRSLLINAVEASCIRTREL 268


RESULT 13
Q8GE04
ID   Q8GE04      PRELIMINARY;      PRT;   274 AA.
AC   Q8GE04;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyrroline-5-carboxylate reductase (EC 1.5.1.2) (Fragment).
OS   Heliobacillus mobilis.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Heliobacteriaceae;
OC   Heliobacillus.
OX   NCBI_TaxID=28064;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22337798; PubMed=12446909;
RA   Raymond J., Zhaxybayeva O., Gogarten J.P., Gerdes S.Y.,
RA   Blankenship R.E.;
RT   "Whole-genome analysis of photosynthetic prokaryotes.";
RL   Science 298:1616-1620(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Liolios K.G., Chu L., Ostrovskaya O., Mendybaeva N., Koukharenko V.,
RA   Gerdes S., Kyrpides N., Overbeek R.;
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY142817; AAN87421.1; -.
KW   Oxidoreductase.
FT   NON_TER     274    274
SQ   SEQUENCE   274 AA;  28410 MW;  7DD2A2EEF5DBDC11 CRC64;

  Query Match             40.1%;  Score 555.5;  DB 2;  Length 274;
  Best Local Similarity   43.4%;  Pred. No. 1.2e-36;
  Matches  115;  Conservative   51;  Mismatches   96;  Indels    3;  Gaps    2;

Qy          9 RRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQA-LGCRTTHSNQEVLQ 67
              ||:||:| | || |:  ||:|    |  | |  | |     | : |: :  :||  |  | :
Db          9 RRIGFIGGGAMAEALLSGLVRL--VPPQAITVSDPNPSRLAYLQSEMAVQTTSDNSVVCE 66

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNL 127
                :|: | || ||||||  :   : ::|:::|:|||::|: ||  | |   ::||:||
Db         67 QSDIVVLAVKPQVLPAVLTPLKTAICSKHLVISIAAGITLAQLESWLSPEVPLIRVMPNT 126

Qy        128 PCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAF 187
              | :: :||  :| || |   : :  : ||||  |  | :|| |::  || ||||  |:|  |
Db        127 PALIGKGASALAGGRAVSEQDVQTAKALLEAVGIAEVLPETYLNAVTGFSGSGPAYVYLF 186

Qy        188 SEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHA 247
               ||| :  || |:   || ::| ||::|:|:::|:  | |||  |  | :||||| |: |
Db        187 IEALIDAGVKEGLTRDLARKLAIQTVIGSAQLLIETGNHPAIERDKVTSPGGTTIAGVQA 246

Qy        248 LEQGGLRAATMSAVEAATCRAKELS 272
              || || |:|  :|| | | ||:|||
Db        247 LEDGGFRSAVFAAVHAGTKRARELS 271


RESULT 14
Q8PYP4
ID   Q8PYP4      PRELIMINARY;      PRT;   270 AA.
AC   Q8PYP4;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Pyrroline-5-carboxylate reductase (EC 1.5.1.2).
GN   MM0817.
OS   Methanosarcina mazei (Methanosarcina frisia).
OC   Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC   Methanosarcinaceae; Methanosarcina.
OX   NCBI_TaxID=2209;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Goe1 / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;
RX   MEDLINE=22120827; PubMed=12125824;
RA   Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA   Martinez-Arias R., Henne A., Wiezer A., Baeumer S., Jacobi C.,
RA   Brueggemann H., Lienard T., Christmann A., Boemecke M., Steckel S.,
RA   Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA   Fritz H.-J., Gottschalk G.;
RT   "The genome of Methanosarcina mazei: evidence for lateral gene
RT   transfer between Bacteria and Archaea.";
RL   J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR   EMBL; AE013306; AAM30513.1; -.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Oxidoreductase; Complete proteome.
SQ   SEQUENCE   270 AA;  28070 MW;  6C7F6327A333B747 CRC64;

  Query Match             37.8%;  Score 524.5;  DB 17;  Length 270;
  Best Local Similarity   43.1%;  Pred. No. 3.5e-34;
  Matches  115;  Conservative   53;  Mismatches   98;  Indels    1;  Gaps    1;

Qy          9 RRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQA-LGCRTTHSNQEVLQ 67
              |::||:|||:|  |: ||:|:|| |: ::| ||   :  |   |  || | :  |  ::: 
Db          4 RKIGFIGAGKMGSALMQGIIKAGIVQPENIGASDVYEPFLKELQTKLGIRVSTDNTVIVR 63

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNL 127
               ::| | ||  | :||  :  :|:| :::|:|||| ||| |  |   |||:||:||: 
Db         64 ESDILILAVKPQTLGSVLENLKNEITSEKLVISIAAGVPLSTYEGALLEGTRVVRVMPNI 123

Qy        128 PCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAF 187
                | | |  :: |::    : |    :  | |   :|||: :|  ||||||| ||: 
Db        124 AATVSEAASGISPGKNATPEDMKDALTIFSAVGTAVQVPESLMDAVTGLSGSGPAFIFPV 183

Qy        188 SEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHA 247
              ||:|:|||  ||    |  :|||||:|| ||| |  | || :|:  | :| |||| |:|| 
Db        184 IEAMADGAVLEGMDRKSALTLAAQTVLGAAKMALETGMHPGELKDMVTSPAGTTIQGVHA 243

Qy        248 LEQGGLRAATMSAVEAATCRAKELSRK 274
              ||: |:||| |:||  |: |::|| :| 
Db        244 LEEAGIRAAFMNAVIRASERSRELGKK 270


RESULT 15
Q9V3F8
ID   Q9V3F8      PRELIMINARY;      PRT;   280 AA.
AC   Q9V3F8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   P5CR protein.
GN   P5CR OR CG6009.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CS;
RA   Misener S.R., Walker V.K.;
RT   "High density of unrelated genes showing overlapping and intraintronic
RT   transcription units in Drosophila.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003724; AAF55626.1; -.
DR   EMBL; AF170829; AAD49740.1; -.
DR   FlyBase; FBgn0015781; P5cr.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Oxidoreductase.
SQ   SEQUENCE   280 AA;  29621 MW;  D3D4FD9611574684 CRC64;

  Query Match             36.5%;  Score 505.5;  DB 5;  Length 280;
  Best Local Similarity   40.9%;  Pred. No. 1.2e-32;
  Matches  112;  Conservative   44;  Mismatches  109;  Indels    9;  Gaps    2;

Qy         10 RVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTHSNQEVLQSC 69
              ::|:| | || ||  ||:| | |:|  :  | |   ||  :: ||  |   |  ||: 
Db          7 KIGFIGGGNMAYAIGSGLVRCGIVKASQVQVSGPHIENLQRWRDLGAVTCDDNCMVLEHS 66

Qy         70 LLVIFATKPHVLPAVLAE-----VAPVVTTEHILVSVAAGVSLSTLEELL----PPNTRV 120
                :|    |||:|    |:      |         ::||| || || ||||          :| 
Db         67 DIVFICVKPHMLTPCAAQLKYKHVPSAKDASKLVVSVLAGTSLETLEEAFSFMGSSELKV 126

Qy        121 LRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSG 180
              :| :||    | ||  |      |   : : :  :| | |  ::|||: :|  ||::| | 
Db        127 IRTMPNTSMQVGEGCTVYTGNARVSHHDLEKIHLMLNALGLAQQVPESMIDAVTGVAGCG 186

Qy        181 VAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGT 240
               |||    ||||:| || |:|  :| : |||||| || :|  |:||| || :||:|||
Db        187 PAFVYTIIEALADGGVKQGVPRQMALQFAAQTLLGAAKTVLLTGKHPAVLRDEVCSPGGA 246

Qy        241 TIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274
              || |:| ||:| ||:  ::||| :: |: || :| 
Db        247 TIVGVHELEKGNLRSTLINAVEKSSQRSAELGKK 280


Search completed: July 30, 2003, 15:25:38
Job time : 102 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 30, 2003, 15:23:05 ; Search time 24 Seconds
(without alignments)
536.888 Million cell updates/sec

Title: US-09-806-536A-14
Perfect score: 1386
Sequence: 1 MAAAEPSPRRVGFVGAGRMA.....AATMSAVEAATCRAKELSRK 274

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	584.5	42.2	319	1	PROC_HUMAN	P32322 homo sapien
2	533.5	38.5	270	1	PROC_METAC	Q9hh99 methanosarc
3	502.5	36.3	273	1	PROC_PEA	Q04708 pisum sativ
4	498	35.9	278	1	PROC_ACTCH	O04016 actinidia c
5	492	35.5	274	1	PROC_SOYBN	P17817 glycine max
6	475	34.3	276	1	PROC_ARATH	P54904 arabidopsis
7	423.5	30.6	269	1	PROC_ECOLI	P00373 escherichia
8	412.5	29.8	267	1	PROC_SYNY3	P74572 synechocyst
9	390.5	28.2	263	1	PROC_TREPA	P27771 treponema p
10	387	27.9	261	1	PROC_THETH	P54893 thermus the
11	383.5	27.7	294	1	PROC_MYCLE	P46725 mycobacteri
12	378	27.3	297	1	PROH_BACSU	P14383 bacillus su
13	366	26.4	295	1	PROC_MYCTU	Q11141 mycobacteri
14	362.5	26.2	270	1	PROC_CORGL	P46540 corynebacte
15	362	26.1	299	1	PROC_CAEEL	Q20848 caenorhabdi
16	345.5	24.9	272	1	PROC_PSEAE	P22008 pseudomonas
17	344	24.8	282	1	PROC_SCHPO	Q9p7y7 schizosacch
18	337.5	24.4	278	1	PROI_BACSU	P54552 bacillus su
19	321	23.2	286	1	PROC_YEAST	P32263 saccharomyc
20	318.5	23.0	265	1	PROC_AQUAE	O66553 aquifex aeo
21	311.5	22.5	275	1	PROC_PASMU	Q9cpe8 pasteurella
22	310.5	22.4	311	1	PROC_NEUCR	Q12641 neurospora
23	307	22.2	278	1	PROC_VIBAL	P52053 vibrio algi
24	305.5	22.0	320	1	PROC_ZALAR	Q12740 zalerion ar
25	301.5	21.8	271	1	PROC_HAEIN	P43869 haemophilus
26	198	14.3	272	1	PROG_BACSU	Q00777 bacillus su
27	189.5	13.7	273	1	CME4_BACSU	P39696 bacillus su
28	164	11.8	257	1	PROC_HELPJ	Q9zk56 helicobacte
29	159	11.5	257	1	PROC_HELPY	O25773 helicobacte
30	125.5	9.1	251	1	PROC_METSM	P22350 methanobrev
31	114	8.2	294	1	GARR_ECOLI	P23523 escherichia
32	99.5	7.2	388	1	ALR_MYCAV	Q91888 mycobacteri
33	98.5	7.1	398	1	ACK1_VIBVU	Q8dah8 vibrio vuln
34	97	7.0	1503	1	MRP6_HUMAN	O95255 homo sapien
35	96	6.9	401	1	CSD_PSEAE	Q9hxx3 pseudomonas
36	96	6.9	528	1	SERA_MYCTU	O53243 mycobacteri
37	94.5	6.8	388	1	ALR_MYCLE	P38056 mycobacteri
38	94	6.8	4391	1	PGBM_HUMAN	P98160 homo sapien
39	93.5	6.7	349	1	GPDA_MYCLE	Q9cbr9 mycobacteri
40	93.5	6.7	398	1	ACK1_VIBPA	Q87mz4 vibrio para
41	93	6.7	528	1	SERA_MYCLE	O33116 mycobacteri
42	93	6.7	2436	1	ABC2_HUMAN	Q9bzc7 homo sapien
43	92.5	6.7	278	1	HBD_DEIRA	Q9rvg1 deinococcus
44	92.5	6.7	564	1	LPP_RAT	O88202 rattus norv
45	92	6.6	386	1	ALR_MYCTU	Q50705 mycobacteri

ALIGNMENTS

SEQUENCE COMPARISON 'A'

RESULT 1
PROC_HUMAN
ID PROC_HUMAN STANDARD; PRT; 319 AA.
AC P32322; Q96DI6;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Pyrroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
GN PYCR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92112821; PubMed=1730675;
RA Dougherty K.M., Brandriss M.C., Valle D.;
RT "Cloning human pyrroline-5-carboxylate reductase cDNA by
RT complementation in Saccharomyces cerevisiae.";
RL J. Biol. Chem. 267:871-875(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyrroline-5-
CC carboxylate + NAD(P)H.
CC -!- PATHWAY: Proline biosynthesis; third (last) step.
CC -!- SUBUNIT: HOMODECAMER OR HOMODODECAMER.
CC -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC FAMILY.
CC ---

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; M77836; AAA36407.1; -.
DR   EMBL; BC001504; AAH01504.1; -.
DR   PIR; A41770; A41770.
DR   Genew; HGNC:9721; PYCR1.
DR   GK; P32322; -.
DR   MIM; 179035; -.
DR   GO; GO:0004735; F:pyrroline 5-carboxylate reductase activity; TAS.
DR   GO; GO:0006561; P:proline biosynthesis; TAS.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Oxidoreductase; Proline biosynthesis; NADP.
FT   CONFLICT    155    155       T -> S (IN REF. 2).
SQ   SEQUENCE   319 AA;  33374 MW;  F5E74B5BDFB475EF CRC64;

  Query Match             42.2%;  Score 584.5;  DB 1;  Length 319;
  Best Local Similarity   46.4%;  Pred. No. 2.3e-37;
  Matches  124;  Conservative   47;  Mismatches   89;  Indels    7;  Gaps    3;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
              |||:|||::| |:|:|   || : |  |:||:| | :|    ||   | : |  |:| :|
Db          3 VGFIGAGQLAFALAKGFTAAGVLAAHKIMASSP-DMDLATVSALRKMGVKLTPHNKETVQ 61

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLP---PNTRVLRVL 124
              ::  | |||::| :| |:   :   ||:|| ||||::|::|: |    |  ||:| :
Db         62 HSDVLFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRPAPRVIRCM 121

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
              | | ||:||| | | | |    : :|:: ||   | | || |  :|  |||||||  |:
Db        122 TNTPVVVREGATVYATGTHAQVEDGRLMEQLLSTVGFCTEVEEDLIDAVTGLSGSGPAYA 181

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
              :|||:| ||||:|  || |: || ||| ||||||  ||| ||: :| :||| ||:
Db        182 FTALDALADGGVKMGLPRRLAVRLGAQALLGAAKMLLHSEQHPGQLKDNVSSPGGATIHA 241

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
              || || || |:  ::|||||:  | :||
Db        242 LHVLESGGFRSLLINAVEASCIRTREL 268


RESULT 2
PROC_METAC
ID   PROC_METAC     STANDARD;      PRT;   270 AA.
AC   Q9HH99;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyrroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
GN   PROC OR MA4102.
OS   Methanosarcina acetivorans.
OC   Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC   Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX   NCBI_TaxID=2214;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C2A / ATCC 35395 / DSM 2834;
RX   MEDLINE=21833514; PubMed=11844777;
RA   Zhang J.K., White A.K., Kuettner H.C., Boccazzi P., Metcalf W.W.;
RT   "Directed mutagenesis and plasmid-based complementation in the
RT   methanogenic archaeon Methanosarcina acetivorans C2A demonstrated by
RT   genetic analysis of proline biosynthesis.";
RT   J. Bacteriol. 184:1449-1454(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C2A / ATCC 35395 / DSM 2834;
RX   MEDLINE=21929760; PubMed=11932238;
RA   [illegible]an J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA   [illegible]h W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA   Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,
RA   Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA   Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA   Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA   Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA   Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA   Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA   Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA   Metcalf W.W., Birren B.;
RT   "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT   and physiological diversity.";
RL   Genome Res. 12:532-542(2002).
CC   -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyrroline-5-
CC       carboxylate + NAD(P)H.
CC   -!- PATHWAY: Proline biosynthesis; third (last) step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC       FAMILY.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; AF305580; AAG22033.1; -.
DR   EMBL; AE011122; AAM07450.1; -.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
SQ   SEQUENCE   270 AA;  27943 MW;  50EC656AFC10B1CF CRC64;

  Query Match             38.5%;  Score 533.5;  DB 1;  Length 270;
  Best Local Similarity   43.4%;  Pred. No. 1.4e-33;
  Matches  116;  Conservative   53;  Mismatches   97;  Indels    1;  Gaps    1;

Qy          9 RRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQA-LGCRTTHSNQEVLQ 67
              :::||:|||:|  |: || |:|| |  ::| ||   :  |   || || | :  |  :::
Db          4 QKIGFIGAGKMGSALMQGTIKAGIVTPENIGASDVYEPFLKDLQAKLGIRVSTDNAVIVR 63

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNL 127
                 ::| | ||  | :||: :   :|:| :::|:|||| ||| |: |   |||:||:||:
Db         64 ESDILILAVKPQTLSSVLSNLKNEITSEKLVISIAAGVPLSTYEDALLEGTRVVRVMPNI 123

Qy        128 PCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAF 187
                | |  :| |::   : |   :  | |   :|||: :|  ||||||| ||:
Db        124 AATVSEAASGIAPGKNATPEDLKAALEIFSAVGTAVQVPESLMDAVTGLSGSGPAFIFPV 183

Qy        188 SEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHA 247
                ||:|:|||  ||    |  :||||:|| ||| |  | || :|:  | :| |||| |:|:
Db        184 IEAMADGAVLEGMDRKSALTLAAQTVLGAAKMALETGMHPGELKDMVTSPAGTTIQGIHS 243

Qy        248 LEQGGLRAATMSAVEAATCRAKELSRK 274
              ||: |:||| |:||  |: |:||| :|
Db        244 LEEAGIRAAFMNAVIRASERSKELGKK 270


RESULT 3
PROC_PEA
ID   PROC_PEA       STANDARD;      PRT;   273 AA.
AC   Q04708;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyrroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
GN   PROC.
OS   Pisum sativum (Garden pea).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae; Pisum.
```

OM protein - protein search, using sw model

Run on: July 30, 2003, 15:23:05 ; Search time 24 Seconds
(without alignments)
536.888 Million cell updates/sec

Title: US-09-806-536A-14
Perfect score: 1386
Sequence: 1 MAAAEPSPRRVGFVGAGRMA.....AATMSAVEAATCRAKELSRK 274

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	584.5	42.2	319	1	PROC_HUMAN	P32322 homo sapien
2	533.5	38.5	270	1	PROC_METAC	Q9hh99 methanosarc
3	502.5	36.3	273	1	PROC_PEA	Q04708 pisum sativ
4	498	35.9	278	1	PROC_ACTCH	O04016 actinidia c
5	492	35.5	274	1	PROC_SOYBN	P17817 glycine max
6	475	34.3	276	1	PROC_ARATH	P54904 arabidopsis
7	423.5	30.6	269	1	PROC_ECOLI	P00373 escherichia
8	412.5	29.8	267	1	PROC_SYNY3	P74572 synechocyst
9	390.5	28.2	263	1	PROC_TREPA	P27771 treponema p
10	387	27.9	261	1	PROC_THETH	P54893 thermus the
11	383.5	27.7	294	1	PROC_MYCLE	P46725 mycobacteri
12	378	27.3	297	1	PROH_BACSU	P14383 bacillus su
13	366	26.4	295	1	PROC_MYCTU	Q11141 mycobacteri
14	362.5	26.2	270	1	PROC_CORGL	P46540 corynebacte
15	362	26.1	299	1	PROC_CAEEL	Q20848 caenorhabdi
16	345.5	24.9	272	1	PROC_PSEAE	P22008 pseudomonas
17	344	24.8	282	1	PROC_SCHPO	Q9p7y7 schizosacch
18	337.5	24.4	278	1	PROI_BACSU	P54552 bacillus su
19	321	23.2	286	1	PROC_YEAST	P32263 saccharomyc
20	318.5	23.0	265	1	PROC_AQUAE	O66553 aquifex aeo
21	311.5	22.5	275	1	PROC_PASMU	Q9cpe8 pasteurella
22	310.5	22.4	311	1	PROC_NEUCR	Q12641 neurospora
23	307	22.2	278	1	PROC_VIBAL	P52053 vibrio algi
24	305.5	22.0	320	1	PROC_ZALAR	Q12740 zalerion ar
25	301.5	21.8	271	1	PROC_HAEIN	P43869 haemophilus
26	198	14.3	272	1	PROG_BACSU	Q00777 bacillus su
27	189.5	13.7	273	1	CME4_BACSU	P39696 bacillus su
28	164	11.8	257	1	PROC_HELPJ	Q9zk56 helicobacte
29	159	11.5	257	1	PROC_HELPY	O25773 helicobacte
30	125.5	9.1	251	1	PROC_METSM	P22350 methanobrev
31	114	8.2	294	1	GARR_ECOLI	P23523 escherichia
32	99.5	7.2	388	1	ALR_MYCAV	Q91888 mycobacteri
33	98.5	7.1	398	1	ACK1_VIBVU	Q8dah8 vibrio vuln
34	97	7.0	1503	1	MRP6_HUMAN	O95255 homo sapien
35	96	6.9	401	1	CSD_PSEAE	Q9hxx3 pseudomonas
36	96	6.9	528	1	SERA_MYCTU	O53243 mycobacteri
37	94.5	6.8	388	1	ALR_MYCLE	P38056 mycobacteri
38	94	6.8	4391	1	PGBM_HUMAN	P98160 homo sapien
39	93.5	6.7	349	1	GPDA_MYCLE	Q9cbr9 mycobacteri
40	93.5	6.7	398	1	ACK1_VIBPA	Q87mz4 vibrio para
41	93	6.7	528	1	SERA_MYCLE	O33116 mycobacteri
42	93	6.7	2436	1	ABC2_HUMAN	Q9bzc7 homo sapien
43	92.5	6.7	278	1	HBD_DEIRA	Q9rvg1 deinococcus
44	92.5	6.7	564	1	LPP_RAT	O88202 rattus norv
45	92	6.6	386	1	ALR_MYCTU	Q50705 mycobacteri

ALIGNMENTS

RESULT 1
PROC_HUMAN
ID PROC_HUMAN STANDARD; PRT; 319 AA.
AC P32322; Q96DI6;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Pyrroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
GN PYCR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92112821; PubMed=1730675;
RA Dougherty K.M., Brandriss M.C., Valle D.;
RT "Cloning human pyrroline-5-carboxylate reductase cDNA by
RT complementation in Saccharomyces cerevisiae.";
RL J. Biol. Chem. 267:871-875(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyrroline-5-
CC carboxylate + NAD(P)H.
CC -!- PATHWAY: Proline biosynthesis; third (last) step.
CC -!- SUBUNIT: HOMODECAMER OR HOMODODECAMER.
CC -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC FAMILY.
CC ---

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M77836; AAA36407.1; -.
DR   EMBL; BC001504; AAH01504.1; -.
DR   PIR; A41770; A41770.
DR   Genew; HGNC:9721; PYCR1.
DR   GK; P32322; -.
DR   MIM; 179035; -.
DR   GO; GO:0004735; F:pyrroline 5-carboxylate reductase activity; TAS.
DR   GO; GO:0006561; P:proline biosynthesis; TAS.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Oxidoreductase; Proline biosynthesis; NADP.
FT   CONFLICT    155    155       T -> S (IN REF. 2).
SQ   SEQUENCE   319 AA;  33374 MW;  F5E74B5BDFB475EF CRC64;

  Query Match             42.2%;  Score 584.5;  DB 1;  Length 319;
  Best Local Similarity   46.4%;  Pred. No. 2.3e-37;
  Matches  124;  Conservative   47;  Mismatches   89;  Indels    7;  Gaps    3;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL---GCRTTHSNQEVLQ 67
              |||:|||::| |:|:|   || : |  |:||:| | :|    ||   | : |  |:| :|
Db          3 VGFIGAGQLAFALAKGFTAAGVLAAHKIMASSP-DMDLATVSALRKMGVKLTPHNKETVQ 61

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLP---PNTRVLRVL 124
                ::  | |||::| :| |:   :   ||:|| ||||::|::|: |    |  ||:| :
Db         62 HSDVLFLAVKPHIIPFILDEIGADIEDRHIVVSCAAGVTISSIEKKLSAFRPAPRVIRCM 121

Qy        125 PNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
               | | ||:||| | | | |    : :|:: ||   | | || |  :|  |||||| |:
Db        122 TNTPVVVREGATVYATGTHAQVEDGRLMEQLLSTVGFCTEVEEDLIDAVTGLSGSGPAYA 181

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
               :|||:| ||||:|  || |: || ||| |||||  ||| ||: :| :||| ||:
Db        182 FTALDALADGGVKMGLPRRLAVRLGAQALLGAAKMLLHSEQHPGQLKDNVSSPGGATIHA 241

Qy        245 LHALEQGGLRAATMSAVEAATCRAKEL 271
              || || || |:  ::|||:  | :||
Db        242 LHVLESGGFRSLLINAVEASCIRTREL 268


RESULT 2
PROC_METAC
ID   PROC_METAC     STANDARD;      PRT;   270 AA.
AC   Q9HH99;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyrroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
GN   PROC OR MA4102.
OS   Methanosarcina acetivorans.
OC   Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC   Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX   NCBI_TaxID=2214;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C2A / ATCC 35395 / DSM 2834;
RX   MEDLINE=21833514; PubMed=11844777;
RA   Zhang J.K., White A.K., Kuettner H.C., Boccazzi P., Metcalf W.W.;
RT   "Directed mutagenesis and plasmid-based complementation in the
RT   methanogenic archaeon Methanosarcina acetivorans C2A demonstrated by
RT   genetic analysis of proline biosynthesis.";
RL   J. Bacteriol. 184:1449-1454(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C2A / ATCC 35395 / DSM 2834;
RX   MEDLINE=21929760; PubMed=11932238;
RA   Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA   FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA   Allen N., Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,
RA   Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA   Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA   Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA   Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA   Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA   Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA   Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA   Metcalf W.W., Birren B.;
RT   "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT   and physiological diversity.";
RL   Genome Res. 12:532-542(2002).
CC   -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyrroline-5-
CC       carboxylate + NAD(P)H.
CC   -!- PATHWAY: Proline biosynthesis; third (last) step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC       FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF305580; AAG22033.1; -.
DR   EMBL; AE011122; AAM07450.1; -.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
SQ   SEQUENCE   270 AA;  27943 MW;  50EC656AFC10B1CF CRC64;

  Query Match             38.5%;  Score 533.5;  DB 1;  Length 270;
  Best Local Similarity   43.4%;  Pred. No. 1.4e-33;
  Matches  116;  Conservative   53;  Mismatches   97;  Indels    1;  Gaps    1;

Qy          9 RRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQA-LGCRTTHSNQEVLQ 67
              :::||:|||:|  |: || |:|| |  ::| ||   :  |   || || | :  |  :::
Db          4 QKIGFIGAGKMGSALMQGTIKAGIVTPENIGASDVYEPFLKDLQAKLGIRVSTDNAVIVR 63

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNL 127
                 ::| | ||  | :||: :   :|:| :::|:|||| ||| |: |   |||:||:||:
Db         64 ESDILILAVKPQTLSSVLSNLKNEITSEKLVISIAAGVPLSTYEDALLEGTRVVRVMPNI 123

Qy        128 PCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAF 187
                | | |  :| |::    : |    :  | |   :|||: :|  |||||||  ||:
Db        124 AATVSEAASGIAPGKNATPEDLKAALEIFSAVGTAVQVPESLMDAVTGLSGSGPAFIFPV 183

Qy        188 SEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHA 247
               ||:|:||| ||    |  :||||:|| ||| |  | || :|:  | :| |||| |:|:
Db        184 IEAMADGAVLEGMDRKSALTLAAQTVLGAAKMALETGMHPGELKDMVTSPAGTTIQGIHS 243

Qy        248 LEQGGLRAATMSAVEAATCRAKELSRK 274
              ||: |:||| |:||  |: |:||| :|
Db        244 LEEAGIRAAFMNAVIRASERSKELGKK 270


RESULT 3
PROC_PEA
ID   PROC_PEA       STANDARD;      PRT;   273 AA.
AC   Q04708;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyrroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
GN   PROC.
OS   Pisum sativum (Garden pea).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae; Pisum.
```

```
OX   NCBI_TaxID=3888;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Williamson C.L., Slocum R.D.;
RT   "Molecular cloning and evidence for osmoregulation of the F1-
RT   pyrroline-5-carboxylate reductase (proC) gene in pea (Pisum sativum
RT   L.).";
RL   Plant Physiol. 100:1464-1470(1992).
CC   -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyrroline-5-
CC       carboxylate + NAD(P)H.
CC   -!- PATHWAY: Proline biosynthesis; third (last) step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X62842; CAA44646.1; -.
DR   PIR; T06477; T06477.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Oxidoreductase; Proline biosynthesis; NADP.
SQ   SEQUENCE   273 AA;  28237 MW;  D719F52935966588 CRC64;

  Query Match             36.3%;  Score 502.5;  DB 1;  Length 273;
  Best Local Similarity   39.7%;  Pred. No. 3.1e-31;
  Matches  104;  Conservative   58;  Mismatches   99;  Indels    1;  Gaps    1;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTHSNQEVLQSCL 70
              :||:|||:|| :||:|  |:| : :  |: :         |:::|     || :|:::
Db         13 LGFIGAGKMAESIAKGASRSGVLPSSRIVTAHSNPSRRAAFESIGITVLSSNDDVVRASN 72

Qy         71 LVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNLPCV 130
              :|:|: || ::  |: :: |::| : :|||||||: |  |:|    : | :||:|| |
Db         73 VVVFSVKPQLVKDVVLKLKPLLTKDKLLVSVAAGIKLKDLQE-WAGHERFIRVMPNTPAA 131

Qy        131 VQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAFSEA 190
              | : | ||: |      :  |: |  : |:  :  : : |  |||||| |::    ||
Db        132 VGQAASVMSLGGAATEEDANLISQLFGSIGKIWKADDKFFDAITGLSGSGPAYIYLAIEA 191

Qy        191 LAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHALEQ 250
              ||:| |  |:|  ||  :|:||:|| | |    |:|| ||: || :||||||  |:| ||:
Db        192 LADGGVAAGLPRDLALSLASQTVLGAASMATLSGKHPGQLKDDVTSPGGTTIAGVHELEK 251

Qy        251 GGLRAATMSAVEAATCRAKELS 272
              || |   |:|| ||  |::|||
Db        252 GGFRGTLMNAVVAAAKRSRELS 273


RESULT 4
PROC_ACTCH
ID   PROC_ACTCH     STANDARD;      PRT;   278 AA.
AC   O04016;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyrroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
OS   Actinidia chinensis (Kiwi) (Yangtao).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Actinidiaceae; Actinidia.
OX   NCBI_TaxID=3625;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Axillary bud;
RA   Walton E.F., Podivinsky E., Wu R.M., Reynolds P.H.S., Young L.W.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyrroline-5-
CC       carboxylate + NAD(P)H.
CC   -!- PATHWAY: Proline biosynthesis; third (last) step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U92287; AAC14482.1; -.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Oxidoreductase; Proline biosynthesis; NADP.
SQ   SEQUENCE   278 AA;  29003 MW;  198B6E934F190236 CRC64;

  Query Match             35.9%;  Score 498;  DB 1;  Length 278;
  Best Local Similarity   38.9%;  Pred. No. 7e-31;
  Matches  110;  Conservative   63;  Mismatches   96;  Indels   14;  Gaps    5;

Qy          1 MAAAE--PSPR---RVGFVGAGRMAGAIAQGLIRAGKVEAQHI----LASAPTDRNLCHF 51
              ||||   | |    ::||:|||:|| :||:|::::| : |  |    | ||  :     |
Db          1 MAAAAFVPIPTDTYKLGFIGAGKMAESIARGVVKSGVLPASRIRTAHLGSARREA----F 56

Qy         52 QALGCRTTHSNQEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLE 111
              :: | :    | :|::   ::||: || ::  |: :: |::: : :|||: ||| |  ||
Db         57 ESFGVKVLDRNDQVVEDSDVIIFSVKPQIVKEVVLQLRPLLSEKQLLVSIVAGVKLKELE 116

Qy        112 ELLPPNTRVLRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVD 171
              :    ::| :||:|| |  | | | ||: |      : :|:  |  | |:  :  |   |
Db        117 D-WAGHSRFIRVMPNTPAAVGEAASVMSLGATATGEDGELITKLFGAIGKIWKADEKLFD 175

Qy        172 IHTGLSGSGVAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLR 231
                |||||| |::    ||||:| |  |:|  ||  :|:||:|| | |:   |:|| ||:
Db        176 AVTGLSGSGPAYIFLAIEALADGGVAAGLPRELALGLASQTVLGAASMVARGGKHPGQLK 235

Qy        232 SDVCTPGGTTIYGLHALEQGGLRAATMSAVEAATCRAKELSRK 274
              || : ||||| |:| ||:|| |    |:|| :|| |::|: ::
Db        236 DDVASAGGTTIAGIHELEKGGFRGTLMNAVVSATKRSQEIFKR 278


RESULT 5
PROC_SOYBN
ID   PROC_SOYBN     STANDARD;      PRT;   274 AA.
AC   P17817;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyrroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
OS   Glycine max (Soybean).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX   NCBI_TaxID=3847;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90340278; PubMed=2199815;
RA   Delauney A.J., Verma D.P.S.;
RT   "A soybean gene encoding delta 1-pyrroline-5-carboxylate reductase
RT   was isolated by functional complementation in Escherichia coli and is
RT   found to be osmoregulated.";
RL   Mol. Gen. Genet. 221:299-305(1990).
CC   -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyrroline-5-
CC       carboxylate + NAD(P)H.
```

```
CC   -!- PATHWAY: Proline biosynthesis; third (last) step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES, BUT MOSTLY IN
CC       NODULES.
CC   -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X16352; CAA34401.1; -.
DR   PIR; S10186; S10186.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Oxidoreductase; Proline biosynthesis; NADP.
SQ   SEQUENCE   274 AA;  28586 MW;  933CFCFDD7598B63 CRC64;

  Query Match             35.5%;  Score 492;  DB 1;  Length 274;
  Best Local Similarity   39.6%;  Pred. No. 2e-30;
  Matches  107;  Conservative   51;  Mismatches   96;  Indels   16;  Gaps    3;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCH--------FQALGCRTTHSN 62
              :||:|||:|| :||:| :|:| :    |       |   |        |:: |     ||
Db         13 LGFIGAGKMAESIARGAVRSGVLPPSRI-------RTAVHFNLARRGAFESFGVTVLPSN 65

Qy         63 QEVLQSCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLR 122
                :|::   :|: : || ::  |:::: |::|   :|||||||  |  |:|    | | :|
Db         66 DDVVRESDVVVLSVKPQLVKDVVSKLTPLLTKHKLLVSVAAGTKLKDLQE-WAGNDRFIR 124

Qy        123 VLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVA 182
              |:|| |  | : | ||: |      :  ::  |  : |:  :  | | |  |||||||
Db        125 VMPNTPAAVGQAASVMSLGGSATEEDGNIIAQLFGSIGKIWKAEEKYFDAITGLSGSGPA 184

Qy        183 FVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTI 242
              :|    ||||:| |  |:|  :|:|:||| | |:   |:|| ||: || :||||||
Db        185 YVYLAIEALADGGVAAGLPRDLSLSLASQTVLGAASMVSQTGKHPGQLKDDVTSPGGTTI 244

Qy        243 YGLHALEQGGLRAATMSAVEAATCRAKELS 272
              |:| || || |   |:|| ||  |::|||
Db        245 TGIHELENGGFRGTLMNAVVAAAKRSRELS 274


RESULT 6
PROC_ARATH
ID   PROC_ARATH     STANDARD;      PRT;   276 AA.
AC   P54904;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyrroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
GN   PROC1 OR AT5G14800 OR T9L3_100.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94294559; PubMed=8022935;
RA   Verbruggen N., Villarroel R., van Montagu M.;
RT   "Osmoregulation of a pyrroline-5-carboxylate reductase gene in
RT   Arabidopsis thaliana.";
RL   Plant Physiol. 103:771-781(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Landsberg erecta;
RA   Verbruggen N., Villarroel R., Hua X., van Montagu M.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=21016721; PubMed=11130714;
RA   Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
RA   Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
RA   Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
RA   Nakazaki N., Naruo K., Okumura S., Shinpo S., Takeuchi C., Wada T.,
RA   Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
RA   Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
RA   Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
RA   Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
RA   Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
RA   Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
RA   Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strowmatt C.,
RA   Wagner-McPherson C., Wollam A., Yoakum M., Bell M., Dedhia N.,
RA   Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
RA   Kirchoff K., Toth K., King L., Bahret A., Miller B., Marra M.,
RA   Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
RA   Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
RA   Entian K.-D., Terryn N., Hartley N., Bent E., Johnson S.,
RA   Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,
RA   Ramsperger U., Wedler H., Balke K., Wedler E., Peters S.,
RA   van Staveren M., Dirkse W., Mooijman P., Klein Lankhorst R.,
RA   Weitzenegger T., Bothe G., Rose M., Hauf J., Berneiser S., Hempel S.,
RA   Feldpausch M., Lamberth S., Villarroel R., Gielen J., Ardiles W.,
RA   Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
RA   Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Fransz P.F.;
RT   "Sequence and analysis of chromosome 5 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 408:823-826(2000).
CC   -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyrroline-5-
CC       carboxylate + NAD(P)H.
CC   -!- PATHWAY: Proline biosynthesis; third (last) step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M76538; AAA61346.1; -.
DR   EMBL; Y08951; CAA70148.1; -.
DR   EMBL; AL391149; CAC01879.1; -.
DR   PIR; JQ2334; JQ2334.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Oxidoreductase; Proline biosynthesis; NADP.
SQ   SEQUENCE   276 AA;  28624 MW;  B577A01C92A3A28B CRC64;

  Query Match             34.3%;  Score 475;  DB 1;  Length 276;
  Best Local Similarity   40.0%;  Pred. No. 3.8e-29;
  Matches  106;  Conservative   54;  Mismatches  103;  Indels    2;  Gaps    2;

Qy         10 RVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCH-FQALGCRTTHSNQEVLQS 68
              :|||:|||:|| :||:|:: :| :    |  :  :: |    |:: |     :::||::
Db         12 KVGFIGAGKMAESIARGVVASGVLPPNRICTAVHSNLNRRDVFESFGVNVFSTSEEVVKE 71

Qy         69 CLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNLP 128
                :|||: || |:   : |:   ::   ||||||||: |: |:|      | :||:|| |
Db         72 SDVVIFSVKPQVVKEAVTELKSKLSKNKILVSVAAGIKLNDLQE-WSGQDRFIRVMPNTP 130

Qy        129 CVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAFS 188
                | | ||: |      :  ::  |  | |:  :  |   |  |||||||  |::
```

```
Db        131 AAVGEAASVMSLGTGATEEDGAIVAMLFGAVGKILKADEKMFDAVTGLSGSGPAYIFLAI 190

Qy        189 EALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHAL 248

Db        191 EALADGGVAAGLPRELALSLASQTVLGAATMVSKTGKHPGVLKDDVTSPGGTTIAGVHEL 250

Qy        249 EQGGLRAATMSAVEAATCRAKELSR 273

Db        251 EKGSFRATLMNAVVAAAKRSRELSQ 275


RESULT 7
PROC_ECOLI
ID   PROC_ECOLI     STANDARD;      PRT;   269 AA.
AC   P00373;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyrroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
GN   PROC OR B0386 OR SF0322.
OS   Escherichia coli, and
OS   Shigella flexneri.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562, 623;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   SPECIES=E.coli;
RX   MEDLINE=83116986; PubMed=6296787;
RA   Deutch A.H., Smith C.J., Rushlow K.E., Kretschmer P.J.;
RT   "Escherichia coli delta 1-pyrroline-5-carboxylate reductase: gene
RT   sequence, protein overproduction and purification.";
RL   Nucleic Acids Res. 10:7701-7714(1982).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=E.coli; STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=E.coli;
RA   Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,
RA   Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H.,
RA   Lin D., Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX   MEDLINE=22272406; PubMed=12384590;
RA   Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA   Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA   Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA   Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA   Yu J.;
RT   "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT   through comparison with genomes of Escherichia coli K12 and O157.";
RL   Nucleic Acids Res. 30:4432-4441(2002).
CC   -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyrroline-5-
CC       carboxylate + NAD(P)H.
CC   -!- PATHWAY: Proline biosynthesis; third (last) step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J01665; AAA86433.1; -.
DR   EMBL; AE000145; AAC73489.1; -.
DR   EMBL; U73857; AAB18110.1; -.
DR   EMBL; AE015066; AAN41981.1; -.
DR   PIR; A00385; RDECC.
DR   EcoGene; EG10769; proC.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
SQ   SEQUENCE   269 AA;  28145 MW;  C5E2CED36BCC508D CRC64;

  Query Match             30.6%;  Score 423.5;  DB 1;  Length 269;
  Best Local Similarity   34.6%;  Pred. No. 3e-25;
  Matches   92;  Conservative   56;  Mismatches  117;  Indels    1;  Gaps    1;

Qy          9 RRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPT-DRNLCHFQALGCRTTHSNQEVLQ 67

Db          3 KKIGFIGCGNMGKAILGGLIASGQVLPGQIWVYTPSPDKVAALHDQFGINAAESAQEVAQ 62

Qy         68 SCLLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRVLPNL 127

Db         63 IADIIFAAVKPGIMIKVLSEITSSLNKDSLVVSIAAGVTLDQLARALGHDRKIIRAMPNT 122

Qy        128 PCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFVCAF 187

Db        123 PALVNAGMTSVTPNALVTPEDTADVLNIFRCFGEAEVIAEPMIHPVVGVSGSSPAYVFMF 182

Qy        188 SEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYGLHA 247

Db        183 IEAMADAAVLGGMPRAQAYKFAAQAVMGSAKMVLETGEHPGALKDMVCSPGGTTIEAVRV 242

Qy        248 LEQGGLRAATMSAVEAATCRAKELSR 273

Db        243 LEEKGFRAAVIEAMTKCMEKSEKLSK 268


RESULT 8
PROC_SYNY3
ID   PROC_SYNY3     STANDARD;      PRT;   267 AA.
AC   P74572;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyrroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
GN   PROC OR SLR0661.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97061201; PubMed=8905231;
RA   Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA   Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA   Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
RA   Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
RA   Yamada M., Yasuda M., Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT   entire genome and assignment of potential protein-coding regions.";
RL   DNA Res. 3:109-136(1996).
CC   -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyrroline-5-
CC       carboxylate + NAD(P)H.
CC   -!- PATHWAY: Proline biosynthesis; third (last) step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D90916; BAA18679.1; -.
DR   PIR; S76767; S76767.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
SQ   SEQUENCE   267 AA;  27883 MW;  FD9C428047F391B0 CRC64;

  Query Match             29.8%;  Score 412.5;  DB 1;  Length 267;
  Best Local Similarity   36.2%;  Pred. No. 2e-24;
  Matches   98;  Conservative   50;  Mismatches  110;  Indels   13;  Gaps    4;

Qy         10 RVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQ-ALGCRTTHSNQEVLQS 68
              ::| :| | || ||   ||       : |:   |       : |      | :  |||
Db          4 QLGIIGGGVMAEAILARLIAEKTYAPEEIIVGEPHGARRDYLQKTYQVRVSPDNQEAANV 63

Qy         69 CLLVIFATKPHVLPAVLAEVA-----PVVTTEHILVSVAAGVSLSTLEELLPPNTRVLRV 123
                ::: | || ||  ||| :|     |:|      :|: ||||| :::  | :  ::|
Db         64 SEVLLLAVKPQVLDRVLASLAGGANRPLV------ISILAGVSLQRIQKGFPDHA-IIRA 116

Qy        124 LPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAF 183
              :|| |  |  |   :|  : |   :    : :  | |   ||||  :|  ||:|||| |:
Db        117 MPNTPATVGAGMTAIAANKMVEPDQLAKAKAIFSAVGNVVEVPENLMDAVTGVSGSGPAY 176

Qy        184 VCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIY 243
              |    ||||:| |  |:| ::| ::| ||:||||:::    :||||::  | :||||||
Db        177 VALMIEALADGGVLAGLPRAIAQKLALQTVLGTAELIKETEEHPAQIKDKVTSPGGTTIA 236

Qy        244 GLHALEQGGLRAATMSAVEAATCRAKELSRK 274
              |:  ||: | |:| : || ||  |::|| :|
Db        237 GVAVLEKMGFRSAIIEAVRAAYRRSQELGKK 267


RESULT 9
PROC_TREPA
ID   PROC_TREPA     STANDARD;      PRT;   263 AA.
AC   P27771; O83775;
DT   01-AUG-1992 (Rel. 23, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyrroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
GN   PROC OR TP0797.
OS   Treponema pallidum.
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX   NCBI_TaxID=160;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Nichols;
RX   MEDLINE=90264287; PubMed=2188947;
RA   Gherardini F.C., Hobbs M.M., Stamm L.V., Bassford P.J. Jr.;
RT   "Complementation of an Escherichia coli proC mutation by a gene cloned
RT   from Treponema pallidum.";
RL   J. Bacteriol. 172:2996-3002(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Nichols;
RX   MEDLINE=98332770; PubMed=9665876;
RA   Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA   Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA   Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA   Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA   McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA   Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA   Venter J.C.;
RT   "Complete genome sequence of Treponema pallidum, the syphilis
RT   spirochete.";
RL   Science 281:375-388(1998).
CC   -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyrroline-5-
CC       carboxylate + NAD(P)H.
CC   -!- PATHWAY: Proline biosynthesis; third (last) step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC       FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M73825; AAA27478.1; -.
DR   EMBL; AE001250; AAC65760.1; -.
DR   PIR; D71281; D71281.
DR   TIGR; TP0797; -.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
FT   CONFLICT      1     65       MNVGFLGFGAMGRALAEGLVHAGALQAAQVYACALNQEKLR
FT                                AQCTSLGIGACASVQELVQKSEWI -> MTWDFWVLEQWDG
FT                                RWQKGWCTQERCKRSSVRLCVKSGKVACAVYIFGHRCLRVS
FT                                SGTGTEKVNGF (IN REF. 1).
FT   CONFLICT     77     93       VLRDRQSFQGKVLISLA -> GTARSPIFQESAISC (IN
FT                                REF. 1).
FT   CONFLICT    256    260       VRAAL -> CRWLS (IN REF. 1).
SQ   SEQUENCE   263 AA;  27645 MW;  ED4AD7C54BAF9D61 CRC64;

  Query Match             28.2%;  Score 390.5;  DB 1;  Length 263;
  Best Local Similarity   38.1%;  Pred. No. 9.2e-23;
  Matches   99;  Conservative   41;  Mismatches  111;  Indels    9;  Gaps    5;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNL-CHFQALGCRTTHSNQEVLQSC 69
              |||:| | |  |:|:||: || ::|  : | |      |     :||     | ||::|
Db          3 VGFLGFGAMGRALAEGLVHAGALQAAQVYACALNQEKLRAQCTSLGIGACASVQELVQKS 62

Qy         70 LLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLPPN----TRVLRVLP 125
                :  | ||  :  || :         :|:|:|||:| :  | |   :     | | :||
Db         63 EWIFLAVKPSQISTVLRDRQSF--QGKVLISLAAGMSCAAYEALFAADPHQGIRHLSLLP 120

Qy        126 NLPCVVQEGAIVMARGRH-VGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAFV 184
              |||| |  | :::|  || :   |   |  :|    : | |  ||  |   ::|   ||
Db        121 NLPCQVARG-VIIAEARHTLHHDEHAALLAVLRTVAQVEVVDTAYFAIAGVIAGCAPAFA 179

Qy        185 CAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIYG 244
                | ||||:  |: |:    |:|:||  | ||| :: | | |||||:  ||:| |:|| |
Db        180 AQFIEALADAGVRYGLARDQAYRLAAHMLEGTAALIQHSGVHPAQLKDRVCSPAGSTIRG 239

Qy        245 LHALEQGGLRAATMSAVEAA 264
              : |||: ||| | : || ||
Db        240 VLALEEQGLRRAVIHAVRAA 259


RESULT 10
PROC_THETH
ID   PROC_THETH     STANDARD;      PRT;   261 AA.
AC   P54893;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyrroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
GN   PROC.
OS   Thermus thermophilus.
OC   Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC   Thermus.
```

```
OX   NCBI_TaxID=274;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HB27;
RX   MEDLINE=94168609; PubMed=8123043;
RA   Hoshino T., Kosuge T., Hidaka Y., Tabata K., Nakahara T.;
RT   "Molecular cloning and sequence analysis of the proC gene encoding
RT   delta 1-pyrroline-5-carboxylate reductase from an extremely
RT   thermophilic eubacterium Thermus thermophilus.";
RL   Biochem. Biophys. Res. Commun. 199:410-417(1994).
CC   -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyrroline-5-
CC       carboxylate + NAD(P)H.
CC   -!- PATHWAY: Proline biosynthesis; third (last) step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D25413; BAA05001.1; -.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Oxidoreductase; Proline biosynthesis; NADP.
SQ   SEQUENCE   261 AA;  27819 MW;  3435C3A8F4299B3E CRC64;

  Query Match          27.9%;  Score 387;  DB 1;  Length 261;
  Best Local Similarity   37.7%;  Pred. No. 1.7e-22;
  Matches  101;  Conservative   47;  Mismatches  102;  Indels   18;  Gaps    9;

Qy         10 RVGFVGAGRMAGAIAQGLIRAG--KVEAQHILASAPTDRNLCHFQALGCR-TTHSNQEVL 66
              |: ||| |:|  :| :| :  |  : |   :|   | :|:    :  | |  | ::  :
Db          2 RLAFVGLGKMGRSILKGALERGFLRPEEVGVLGRTP-ERSRELAEPFGVRPLTRADLGMA 60

Qy         67 QSCLLVIFATKPHVLPAVLAEVAPVVTTEHIL--VSVAAGVSLSTLEELLPPNTRVLRVL 124
              :   |: | :|   ||:  |:|      | |  :|: ||:| | |   |  | ||:| :
Db         61 ER---VLIAVQPRDFPALAPEIA-----HHRLGYISIMAGISTSVLARRL-DNRRVVRAM 111

Qy        125 PNLPCVVQEGAIVMARGRHVGSSE-TKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAF 183
              |||  |: | :  :    :    :|     : |    |   |:||   |  ||:| |  |:
Db        112 PNLAVVIGESSTALTALKEAREAEDLAFARALFATVGDVYEIPEHLFDAFTGMSASAPAY 171

Qy        184 VCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIY 243
              :   :|||:  ||||||  :|| |:||  |  | ::|  :|:||||:: :| :||||||:
Db        172 LAVVAEALADAGVKMGMPRALALRLAADALAATGELL--KGRHPAQVKDEVASPGGTTIH 229

Qy        244 GLHALEQGGLRAATMSAVEAATCRAKEL 271
              ||||||   :|||  ||||||  |  ||
Db        230 GLHALEARAVRAAFYEAVEAATRRGHEL 257


RESULT 11
PROC_MYCLE
ID   PROC_MYCLE     STANDARD;      PRT;   294 AA.
AC   P46725;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyrroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
GN   PROC OR ML2430 OR B2168_C2_211.
OS   Mycobacterium leprae.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1769;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Smith D.R., Robison K.;
RL   Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TN;
RX   MEDLINE=21128732; PubMed=11234002;
RA   Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA   Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA   Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA   Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA   Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA   Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA   Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA   Barrell B.G.;
RT   "Massive gene decay in the leprosy bacillus.";
RL   Nature 409:1007-1011(2001).
CC   -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyrroline-5-
CC       carboxylate + NAD(P)H.
CC   -!- PATHWAY: Proline biosynthesis; third (last) step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U00018; AAA17233.1; -.
DR   EMBL; AL583925; CAC31947.1; -.
DR   PIR; S72897; S72897.
DR   Leproma; ML2430; -.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
SQ   SEQUENCE   294 AA;  30237 MW;  EA8606C9CBBB6D9D CRC64;

  Query Match          27.7%;  Score 383.5;  DB 1;  Length 294;
  Best Local Similarity   32.8%;  Pred. No. 3.5e-22;
  Matches   94;  Conservative   57;  Mismatches  113;  Indels   23;  Gaps    4;

Qy          7 SPRRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTHSNQEVL 66
              |  |:  :| | :  |:  ||:|||:     ::|    ||            | |  : :
Db          4 SMARIAIIGGGSIGEALLSGLLRAGRQVKDLVVAERMPDRARYLADTYSVLVT-SVTDAV 62

Qy         67 QSCLLVIFATKPHVLPAVLAEVAPVV-----TTEHILVSVAAGVSLSTLEELLPPNTRVL 121
              :: : |: | ||  : :|: ::         : | :||:|||||::: ||  ||  | |:
Db         63 ENAMFVVVAVKPTDVESVMGDLVQAAAVANDSAEQVLVTVAAGVTITYLESKLPAGTPVV 122

Qy        122 RVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGV 181
              | :||   :|  |  |:|:|| |   : : :  : :| |    |||: :|  | :||||
Db        123 RAMPNAAALVGAGVTVLAKGRFVTGQQFEDVLAMFDAVGGVLTVPESQMDAVTAVSGSGP 182

Qy        182 AFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLH---EGQH---------- 226
              |:     ||| :  | :|:   :|  :||||: |:| |||    : :|
Db        183 AYFFLLVEALVDAGVAVGLTRQVATELAAQTMAGSAAMLLERMDQDRHSAEVAPLGAQVD 242

Qy        227 --PAQLRSDVCTPGGTTIYGLHALEQGGLRAATMSAVEAATCRAKEL 271
                ||||: : :|||||   |  ||:||||    :||:||  |:::|
Db        243 VPAAQLRATITSPGGTTAAALRELERGGLRMVVDAAVQAAKIRSEQL 289


RESULT 12
PROH_BACSU
ID   PROH_BACSU     STANDARD;      PRT;   297 AA.
AC   P14383; O07508; O31828;
DT   01-JAN-1990 (Rel. 13, Created)
```

```
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyrroline-5-carboxylate reductase 1 (EC 1.5.1.2) (P5CR 1) (P5C
DE   reductase 1).
GN   PROH OR PROC.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
RN   [2]
RP   SEQUENCE OF 1-256 FROM N.A.
RC   STRAIN=W23, and 168;
RX   MEDLINE=91192601; PubMed=1849493;
RA   Ahn K.S., Wake R.G.;
RT   "Variations and coding features of the sequence spanning the
RT   replication terminus of Bacillus subtilis 168 and W23 chromosomes.";
RL   Gene 98:107-112(1991).
RN   [3]
RP   SEQUENCE OF 1-200 FROM N.A.
RC   STRAIN=W23;
RX   MEDLINE=89155440; PubMed=2493444;
RA   Lewis P.J., Wake R.G.;
RT   "DNA and protein sequence conservation at the replication terminus in
RT   Bacillus subtilis 168 and W23.";
RL   J. Bacteriol. 171:1402-1408(1989).
RN   [4]
RP   SEQUENCE OF 199-297 FROM N.A., AND FUNCTION.
RC   STRAIN=168 / SMY;
RX   MEDLINE=21311767; PubMed=11418582;
RA   Belitsky B.R., Brill J., Bremer E., Sonenshein A.L.;
RT   "Multiple genes for the last step of proline biosynthesis in Bacillus
RT   subtilis.";
RL   J. Bacteriol. 183:4389-4392(2001).
CC   -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyrroline-5-
CC       carboxylate + NAD(P)H.
CC   -!- PATHWAY: Proline biosynthesis; third (last) step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z99114; CAB13741.1; -.
DR   EMBL; M24523; AAA22722.1; -.
DR   EMBL; AF006720; AAB62697.1; -.
DR   PIR; E69682; E69682.
DR   SubtiList; BG11049; proH.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
FT   VARIANT      58     58       T -> A (IN STRAIN W23).
FT   VARIANT      70     70       L -> S (IN STRAIN W23).
FT   VARIANT      74     74       L -> I (IN STRAIN W23).
FT   VARIANT      93     93       N -> S (IN STRAIN W23).
FT   VARIANT     100    100       S -> T (IN STRAIN W23).
FT   VARIANT     127    127       E -> Q (IN STRAIN W23).
FT   VARIANT     159    159       K -> Q (IN STRAIN W23).
FT   VARIANT     236    236       H -> Q (IN STRAIN W23).
FT   VARIANT     239    239       I -> V (IN STRAIN W23).
FT   CONFLICT    256    256       Q -> E (IN REF. 2).
FT   CONFLICT    264    297       GEAISQAIKHAAKRSKEISEDIEKTAAPLSGVIK -> EKR
FT                                FLKQ (IN REF. 1).
SQ   SEQUENCE   297 AA;  32031 MW;  F4DC7E6F84C6ED42 CRC64;

  Query Match             27.3%;  Score 378;  DB 1;  Length 297;
  Best Local Similarity   32.2%;  Pred. No. 9.2e-22;
  Matches   91;  Conservative   50;  Mismatches  106;  Indels   36;  Gaps    3;

Qy          9 RRVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTHSNQEVL-- 66
              ::| |:||| ||  :  |::|| |:  |:|                  | |  || | |
Db         17 KKVAFIGAGSMAEGMISGIVRANKIPKQNI----------------CVTNRSNTERLTE 59

Qy         67 ------------QSCL----LVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLST 109
                        | |:    ::| | ||      |: :   :    :::|| ||:: |
Db         60 LELQYGIKGALPNQLCIEDMDVLILAMKPKDAENALSSLKSRIQPHQLILSVLAGITTSF 119

Qy        110 LEELLPPNTRVLRVLPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAY 169
              :|: |     |:||:||   ::   |  :| |::|     || : ||   |    : |
Db        120 IEQSLLNEQPVVRVMPNTSSMIGASATAIALGKYVSEDLKKLAEALLGCMGEVYTIQENQ 179

Qy        170 VDIHTGLSGSGVAFVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQ 229
              :|| ||::||| |:     | : :   : |:   |:  | |||||| |||:  |:||
Db        180 MDIFTGIAGSGPAYFYYLMEFIEKTGEEAGLDKQLSRSIGAQTLLGAAKMLMETGEHPEI 239

Qy        230 LRSDVCTPGGTTIYGLHALEQGGLRAATMSAVEAATCRAKELS 272
              || :: :| |||  || ||:: |   |   |:: |  |:||:|
Db        240 LRDNITSPNGTTAAGLQALKKSGGGEAISQAIKHAAKRSKEIS 282


RESULT 13
PROC_MYCTU
ID   PROC_MYCTU     STANDARD;      PRT:   295 AA.
AC   Q11141;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyrroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
GN   PROC OR RV0500 OR MT0520 OR MTCY20G9.26.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=H37Rv;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyrroline-5-
CC       carboxylate + NAD(P)H.
CC   -!- PATHWAY: Proline biosynthesis; third (last) step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC       FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Z77162; CAB00926.1; -.
DR   EMBL; AE006952; AAK44743.1; -.
DR   PIR; G70745; G70745.
DR   TIGR; MT0520; -.
DR   TubercuList; Rv0500; -.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
SQ   SEQUENCE   295 AA;  30171 MW;  F95C3F407BE5408F CRC64;

  Query Match             26.4%;  Score 366;  DB 1;  Length 295;
  Best Local Similarity   31.6%;  Pred. No. 7.4e-21;
  Matches   90;  Conservative   55;  Mismatches  116;  Indels   24;  Gaps    3;

Qy         10 RVGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTHSNQEVLQSC 69
              |:  :| | :  |:  ||:|||:     ::|    ||     |      | |  : :::
Db          7 RIAIIGGGSIGEALLSGLLRAGRQVKDLVVAERMPDRANYLAQTYSVLVT-SAADAVENA 65

Qy         70 LLVIFATKPHVLPAVLAEVAPVV------TTEHILVSVAAGVSLSTLEELLPPNTRVLRV 123
                |: | ||  :  |:|::|         : | : |:| ||:::  |  ||  | |:|
Db         66 TFVVVAVKPADVEPVIADLANATAAAENDSAEQVFVTVVAGITIAYFESKLPAGTPVVRA 125

Qy        124 LPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAF 183
               :||   :|  |   :|:|| |   : : :  | :| |    |||: :|  | :|||| |:
Db        126 MPNAAALVGAGVTALAKGRFVTPQQLEEVSALFDAVGGVLTVPESQLDAVTAVSGSGPAY 185

Qy        184 VCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQH----------------- 226
                   ||| :  | :|:   :|  :|||| |:| |||   :
Db        186 FFLLVEALVDAGVGVGLSRQVATDLAAQTMAGSAAMLLERMEQDQGGANGELMGLRVDLT 245

Qy        227 PAQLRSDVCTPGGTTIYGLHALEQGGLRAATMSAVEAATCRAKEL 271
                ::||: | :|||||   |  ||:|| | |  :||:||  |:::|
```

```
Db        246 ASRLRAAVTSPGGTTAAALRELERGGFRMAVDAAVQAAKSRSEQL 290


RESULT 14
PROC_CORGL
ID   PROC_CORGL     STANDARD;      PRT;   270 AA.
AC   P46540;
DT   01-NOV-1995 (Rel. 32, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Pyrroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).
GN   PROC OR CGL0410.
OS   Corynebacterium glutamicum (Brevibacterium flavum).
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX   NCBI_TaxID=1718;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 17965 / Melassecola;
RX   MEDLINE=96345604; PubMed=8755867;
RA   Ankri S., Serebrijski I., Reyes O., Leblon G.;
RT   "Mutations in the Corynebacterium glutamicum proline biosynthetic
RT   pathway: a natural bypass of the proA step.";
RL   J. Bacteriol. 178:4412-4419(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA   Nakagawa S.;
RT   "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyrroline-5-
CC       carboxylate + NAD(P)H.
CC   -!- PATHWAY: Proline biosynthesis; third (last) step.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC       FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U31225; AAC44172.1; -.
DR   EMBL; AP005275; BAB97803.1; -.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
FT   CONFLICT    270    270       R -> H (IN REF. 1).
SQ   SEQUENCE   270 AA;  28242 MW;  85F50E233D94158F CRC64;

  Query Match             26.2%;  Score 362.5;  DB 1;  Length 270;
  Best Local Similarity   30.6%;  Pred. No. 1.2e-20;
  Matches   83;  Conservative   63;  Mismatches  114;  Indels   11;  Gaps    4;

Qy         11 VGFVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQAL----GCRTTHSNQEVL 66
              :  :| |::  |:  ||| |  :  |:|     |:|:    | |    |       | :
Db          4 IAVIGGGQIGEALVSGLI-AANMNPQNIRV---TNRSEERGQELRDRYGILNMTDNSQAA 59

Qy         67 QSCLLVIFATKPHVLPAVLAEVAPVV---TTEHILVSVAAGVSLSTLEELLPPNTRVLRV 123
                 :|    ||  :  ||:|:   :   : : ::||:|||:|:: :||       |:||
Db         60 DEADVVFLCVKPKFIVEVLSEITGTLDNNSAQSVVVSMAAGISIAAMEESASAGLPVVRV 119

Qy        124 LPNLPCVVQEGAIVMARGRHVGSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVAF 183
               :|| | :| :|   : :||:| : : : :: ||   |   || |: :|  | :|||  |:
Db        120 MPNTPMLVGKGMSTVTKGRYVDAEQLEQVKDLLSTVGDVLEVAESDIDAVTAMSGSSPAY 179

Qy        184 VCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLHEGQHPAQLRSDVCTPGGTTIY 243
               :  :||| |  | :|:| : | ::|  :  | | |:   |: |::||: | :| |||:
```

```
Db        180 LFLVTEALIEAGVNLGLPRATAKKLAVASFEGAATMMKETGKEPSELRAGVSSPAGTTVA 239

Qy        244 GLHALEQGGLRAATMSAVEAATCRAKELSRK 274
              :  ||: |:| |   | :|   |::|| ::
Db        240 AIRELEESGIRGAFYRAAQACADRSEELGKR 270


RESULT 15
PROC_CAEEL
ID   PROC_CAEEL     STANDARD;      PRT;   299 AA.
AC   Q20848;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Putative pyrroline-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C
DE   reductase).
GN   F55G1.9.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Murray J., Le T.T.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = 1-pyrroline-5-
CC       carboxylate + NAD(P)H.
CC   -!- PATHWAY: Proline biosynthesis; third (last) step.
CC   -!- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U58750; AAB00645.1; -.
DR   PIR; T29226; T29226.
DR   WormPep; F55G1.9; CE07286.
DR   InterPro; IPR000304; P5CR.
DR   Pfam; PF01089; P5CR; 1.
DR   TIGRFAMs; TIGR00112; proC; 1.
DR   PROSITE; PS00521; P5CR; 1.
KW   Hypothetical protein; Oxidoreductase; Proline biosynthesis; NADP.
SQ   SEQUENCE   299 AA;  32057 MW;  75D2D6D32E0A5670 CRC64;

  Query Match             26.1%;  Score 362;  DB 1;  Length 299;
  Best Local Similarity   33.5%;  Pred. No. 1.5e-20;
  Matches   93;  Conservative   53;  Mismatches  114;  Indels   18;  Gaps    6;

Qy         13 FVGAGRMAGAIAQGLIRAGKVEAQHILASAPTDRNLCHFQALGCRTTHSNQ-EVLQ--SC 69
              |:| | || || :|    |     :|:    |:::   :: || :   :|  |:|:  |
Db         19 FIGGGNMAAAIIKGCQNKGFTPKSNIVIGVQTEKSAEKWRQLGYKNVFTNTLEMLERYST 78

Qy         70 LLVIFATKPHVLPAVLAEVAPVVTTEHILVSVAAGVSLSTLEELLP---PNTRVLRVLPN 126
               : :   || |   |::   || :    ::|| ||| |  |   ||    || ::|::||
Db         79 AIYVICVKPQVFEEVVSS-WPVNSRPEFIISVMAGVPLKVLNAKLPFVSGNTTIVRLMPN 137

Qy        127 LPCVVQEGAIVMARGRHV----GSSETKLLQHLLEACGRCEEVPEAYVDIHTGLSGSGVA 182
              :   : ||  |   ::       |  :| :   |  |  | :||   :    : ||  |
Db        138 VASSIGAGASTMCYEKNEKIMNQDSHIELAREFAECVGTVELIPERCFNPAMAIGGSSPA 197

Qy        183 FVCAFSEALAEGAVKMGMPSSLAHRIAAQTLLGTAKMLLH-------EGQHPAQLRSDVC 235
              :   : |:||:|||  |:  : | |:||| :|| |:|:|:       | ||   |:  ||
Db        198 WTFMYIESLADGAVAQGLGRAEAKRLAAQAVLGAAQMVLNSNSGFDIETQHFGSLKDMVC 257

Qy        236 TPGGTTIYGLHALEQGGLRAATMSAVEAATCRAKELSR 273
              :|||||| |: |||: | | | | || ||: :| |:::
Db        258 SPGGTTIEGVRALEKNGFRYAVMEAVVAASTKADEMAK 295
```

```
Search completed: July 30, 2003, 15:23:51
Job time : 32 secs
```